Understanding the Y chromosome variation by haplogroup and haplotype analyses in a Korean population

Myung Jin Park

Department of Medical Science

The Graduate School, Yonsei University

Understanding the Y chromosome variation by haplogroup and haplotype analyses in a Korean population

Directed by Professor Kyoung-Jin Shin

The Doctoral Dissertation
submitted to the Department of Medical Science,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy of Medical Science

Myung Jin Park

December 2012

This certifies that the Doctoral Dissertation of Myung Jin Park is approved.

Thesis Supervisor: Kyoung-Jin Shin

Thesis Committee Member#1 Yang, Woo Ick

Thesis Committee Member#2 Lee, Hwan Young

Thesis Committee Member#3, Lee, Jin Sung

Thesis Committee Member#4, Choi, Jong Rak

The Graduate School
Yonsei University

December 2012

ACKNOWLEDGEMENTS

처음 설레는 마음으로 법의학과에 들어와 여러 선생님들을 만나고 연구실에서 공부하던 지난날들이 생각납니다. 쉽지 않은 결정이었지만 의욕적으로 시작한 박사학위과정이 이제는 끝을 향해 가고 있습니다. 그 동안 웃고 울고 힘들어 하며 함께했던 세월이 쌓이고 쌓여 지금은 너무나도 값진 추억이 되었습니다.

부족한 첫 박사과정 제자를 위해 항상 격려해 주시고 정성으로 지도해 주신 신경진 교수님, 법의학과에 발을 디딜 때부터 지금까지 함께하며 많은 가르침과 대학원생활에 큰 도움을 주신 이환영 교수님, 항상 든든한 버팀목이 되어주신 양우익 교수님, 바쁘신 와중에도 시간을 내어 논문 심사와 지도를 해 주신 이진성 교수님과 최종락 교수님께 감사의 말씀 올립니다.

또한 정신적 지주가 되어주시고 공부할 수 있도록 용기를 주신 박경숙 교수님, 대학원 생활 전 과정에 귀감이 되어 주신 이숭덕 교수님, 항상 멀리 계셔도 응원해 주시는 최종훈, 조상호 교수님께도 감사의 말씀을 드리고 싶습니다.

함께 연구실 생활을 시작한 동기이자 후배가 된 정욱희 선생님, 대학교 후배이자 같이 조교 생활한 김나영 선생님, 늘 귀찮은 일도 마다하지 않고 잘 따라준 아진씨, 후임으로 와서 고생하고 있는 이은영, 안자현 선생님, 그리고 조용히 응원해 주신 양인석 박사님에게도 고마움을 전합니다.

학위를 끝마칠 수 있도록 도와주신 한면수 센터장님, 항상 애정 어린 관심을 주신 박기원 센터장님, 마지막까지 학위 일정을 마칠 수 있도록 배려해 주신 강필원 실장님, 묵묵히 도와주신 김남예 실장님, 많이 아껴주시는 이양한, 조남수 실장님, 격려를 아끼지 않으시던 최동호 팀장님과 여러 실장님들을 비롯한 선생님들에게도 깊은 감사의 마음을 표합니다. 그리고 무사히 학위과정을 끝마칠 수 있도록 많은 관심과 애정을 준 동기들과 제 소중한 친구에게도 감사인사 전합니다.

마지막으로 지금의 제가 있기까지 가장 큰 힘이 되어 준 가족들, 조용히 응원해 주신 아버지, 늘 사랑하는 어머니께 감사 드리고 항상 아껴주는 동생 부부와 귀염둥이 조카에게도 고마움을 전합니다.

저에게 힘이 되어 주신 분 들이 있기에 이제는 세상을 헤쳐 나갈 수 있는 용기가 생겼습니다. 대학원에서 보고 듣고 느낀 것들을 디딤돌 삼아 이 세상에 소금과 같은 존재가 되겠습니다. 모두 감사합니다.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

<ABSTRACT>

Understanding the Y chromosome variation by haplogroup and haplotype analyses in a Korean population

Myung Jin Park

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Kyoung-Jin Shin)

Genetic variation on the non-recombining portion of the Y-chromosome contains information about the ancestral and geographical origin of biological samples as well as differentiation between male lineages. A molecular characterization of the Y-chromosome genetic structure was performed using a combination of Y-chromosomal single nucleotide polymorphisms (Y-SNPs) and Y-chromosomal short tandem repeats (Y-STRs) in a Korean population. Six multiplex PCRs followed by SBE reaction and two multiplex allele-specific PCR assays were developed for the identification of haplogroups in Koreans as well as haplogroups frequent in East Asians. Validation experiments demonstrated that the multiplex PCR systems followed by SBE reaction were optimized for analyzing low template and highly degraded DNA whereas a multiplex allele-specific PCR assay would be useful for simple and reliable determination of haplogroups in a large number of samples. In a test using DNA samples from Korean males, a total of 21 different haplogroups were identified by 33 Y-SNPs including the newly redefined markers (PK4, KL2 and P164) in

haplogroup O. When genotyping the SNPs, phylogenetic nonequivalence was found between SNPs M117 and M133, which define haplogroup O3a2c1a, suggesting that the position of the M133 marker should be changed. This study has shown that the haplotypes consisting of DYS392, DYS393, DYS437, DYS438, DYS448 and DYS388 loci can preserve phylogenetic information by their relatively low mutation rates, and hence can be used to roughly distinguish Y-chromosome haplogroups, whereas more rapidly mutating Y-STRs such as DYS449 and DYS458 are useful for differentiating male lineages. At the relatively rapidly mutating DYS447, DYS449, DYS458 and DYS464 loci, unusually short alleles and intermediate alleles with common sequence structures are informative for elucidating the substructure within the context of a particular haplogroup. In addition, some deletion mutations in the DYS385 flanking region and the null allele at DYS448 were associated with a single haplogroup background. These variants support the hypothesis that the variant originated from a single mutational event like binary markers. The high-resolution haplogroup and haplotype data will improve our understanding of the Korean population substructure and will also help to infer haplogroup background or common ancestry.

Key words : Y chromosome, Haplogroup, Single nucleotide polymorphism, Haplotype, Short tandem repeat, Atypical allele, Korean

Understanding the Y chromosome variation by haplogroup and haplotype analyses in a Korean population

Myung Jin Park

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Kyoung-Jin Shin)

I. Introduction

The paternally inherited non-recombining portion of the Y chromosome (NRY) is targeted in forensic and medical genetics, human evolutionary studies, and genealogical reconstruction.[1-6] Two classes of markers on the NRY, single nucleotide polymorphism (SNP) and short tandem repeat (STR), are widely used. Due to their low mutation rate, Y chromosome SNPs are useful genetic markers for reconstructing male lineages through hierarchically arranged allelic sets known as haplogroups, whereas the high mutation rates of Y-chromosomal STRs make them valuable for differentiating between unrelated males and inferring affinities among related populations.[4-8]

Y-chromosomal haplogroups are distributed non-randomly in populations and geographical regions, allowing us to infer the origin evolution and human history by tracking back male patterns of migration from a modern human population.[5, 8, 9] The analysis of Y-SNPs is of interest in forensic investigation,

because it can show regional specificity, providing useful information about the geographic origin of a subject or evidence.[9, 10] In casework, where there are decomposed bodies or human remains, or victims of mass disasters, including airplane crashes, tsunamis or terrorist attack involving people from various geographical area, the prediction of population origin may be a powerful aid in forensic investigations.[11, 12] In addition, the particular advantage of SNPs is that genotyping can be done in much smaller amplicons than those used in conventional STR analysis and thus, even with highly degraded DNA, where STR typing fails.[12, 13]

Y-chromosomal STRs provide a resolution that is useful as a forensic tool for discrimination between males and identification of male profiles in a mixed sample with female.[1, 14, 15] Y-STR loci are not independent of one another and are co-inherited as extended haplotypes of linked markers. Haplotype frequencies are required to provide a statistical estimate of the significance of a match. The estimation of the frequency of occurrence of a particular Y-STR haplotype needs large haplotype databases. There are several regional Y-STR haplotype databases currently in use globally, such as the Y-Chromosome Haplotype Reference Database (YHRD).[16-18] However, the geographic clustering of haplogroups can lead to large differences in Y-STR haplotype frequencies between different genetic populations, which may have an impact on database composition in a region or a nation.[18] Namely, some Y-STR variability is highly partitioned by differences among haplogroups, which can

affect the level of resolution for the paternal lineage differentiation of the given Y-STR haplotypes.[19, 20] Otherwise, it is indicated that haplogroup prediction can be done from the Y-STR haplotype information.[21, 22] As such, YHRD is complements Y-STR data with information on Y-SNPs[16], and the number of studies integrating Y-SNP and Y-STR data is growing.[23-25] In addition, studies of the relationship between haplogroups and Y-STR variants, such as the relationship of haplogroup affiliations for partial deletion/insertion mutations or intermediate alleles in Y-STRs, have been reported.[25-28]

In Korea, an effort has been made to establish a high-quality Y-STR database. How many and which Y-STRs should be typed and stored is an issue for discussion by forensic geneticists in order to yield high-resolution lineage differentiation. In addition, as the demand for inferring geographic origin is increasing because of crimes involving foreigners and immigrants from other Asians groups, the need for Y-chromosomal haplogroup determination is increasing for forensic DNA analysis. Therefore, a combined analysis of Y-SNP and Y-STR variations is needed to provide basic information on the Korean population.

Meanwhile, the Y-chromosomal haplogroup tree has been periodically updated. In particular, the phylogeny of haplogroup O-M175 has recently been revised to include the phylogenetic positions of redefined markers, L127, KL1, KL2, P164 and PK4.[29, 30] A previous study showed that L127/KL1/KL2 and P164 are highly informative for separating substantial samples belonging to

haplogroup O3a-M324 in the Han Chinese population.[30] The haplogroup O-M175 is one of the major clades in the Korean population and the haplogroup O3a-M324 accounts for 43.9% of Korean males.[31, 32] It is necessary to confirm the relative positions of the redefined markers and to determine the distribution of subhaplogroups in the Korean population.

In this study, to provide the distributions of Korean haplogroup and to analyze the relationship of haplogroup with Y-STR haplotype, multiplex allele-specific PCR assays were developed for the simultaneous detection of Y-SNP genotypes in a hierarchical order and to classify the Korean haplotypes into their corresponding haplogroups, especially with in-depth resolution for haplogroup O-M175 according to the revised phylogenetic tree. In addition, multiplex single base extension (SBE) reactions were developed for scoring Y-SNPs of East Asian haplogroups following a small size amplicon strategy that is suitable for analysis of degraded DNA or ancient DNA analysis (e.g., identification of Korean War and Vietnam War victims, and genetic characterization of ancient remains). Molecular characterizations were performed for 22 Y-STR markers and their haplotypes (DYS19, DYS385a/b, DYS388, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS446, DYS447, DYS448, DYS449, DYS456, DYS458, DYS464, DYS635 and GATA H4.1), which consist of the European minimal (DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392 and DYS393), the Scientific Working Group on DNA Analysis Methods (SWGDAM)

haplotype (minimal haplotype including DYS438 and DYS439), AmpFlSTR® Yfiler™ haplotype loci (SWGDAM haplotype including DYS437, DYS448, DYS456, DYS458, DYS635 and GATA H4.1), and 5 additional Y-STRs (DYS388, DYS446, DYS447, DYS449 and DYS464) that showed high diversities.[33, 34] Finally, the haplogroup affiliation for each usual and variant allele was evaluated to elucidate their relationship with the binary haplogroup and to subsequently identify Y-STR alleles potentially representing a substructure within the haplogroup tree.

II. MATERIAL AND METHODS

1. Y-SNP analysis

A. DNA samples

This study protocol was approved by the Institutional Review Board of Severance Hospital, Yonsei University in Seoul, Korea. DNA samples from 300 unrelated Korean males were obtained from the National Biobank of Korea for validation of the developed multiplex PCR systems. I analyzed DNA samples from 1006 unrelated Korean males, which include 706 individuals who have been typed for 19 or 22 Y-STRs.[33, 34] DNA concentrations were measured using a NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and the final sample concentrations were adjusted to 1.0 ng/μl. For the sensitivity test, 9948 male DNA (Promega, Madison, WI, USA)

and 2800M Control DNA (Promega) was serially diluted to concentrations of 1000, 500, 250, 125, 62, 31, and 15 pg/μl. Artificially degraded DNA was prepared by digesting 1.2 μg of human genomic DNA with 0.02 U of DNase I (NEB, Ipswich, MA, USA) at 37°C for 40 min. DNA degradation to fragment sizes of around 100 bp was confirmed by ethidium bromide staining after agarose gel electrophoresis. Ten DNA samples that were extracted from 55-year-old skeletal remains during a previous study[35] were analyzed to evaluate multiplex SBE reactions. Concentrations of DNA obtained from skeletal remains were determined using the Quantifiler™ Human DNA Quantification Kit and 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA).[35]

B. Multiplex PCR followed by SBE reaction

(A) Y-SNP selection and primer design for PCR amplification and SBE reaction

A set of 27 biallelic Y chromosome markers (M7, M9, M48, M89, M95, M119, M122, M133, M134, M145, M159, M162, M175, M214, M217, M324, M407, P31, P53.1, P101, P164, P201, JST002611, KL2, $RPS4Y_{711}$ (M130), SRY_{465} (M176), and 47z) was selected to determine the world-wide major haplogroups, subhaplogroups O, and subhaplogroups C that are present in East Asian populations, including Koreans.[36-38] The 27 Y-SNPs and the haplogroup tree defined by these markers are shown in Figure 1. The nomenclature and

topology of the Y chromosome haplogroups followed those of Karafet et al.[29] and Yan et al.[30]

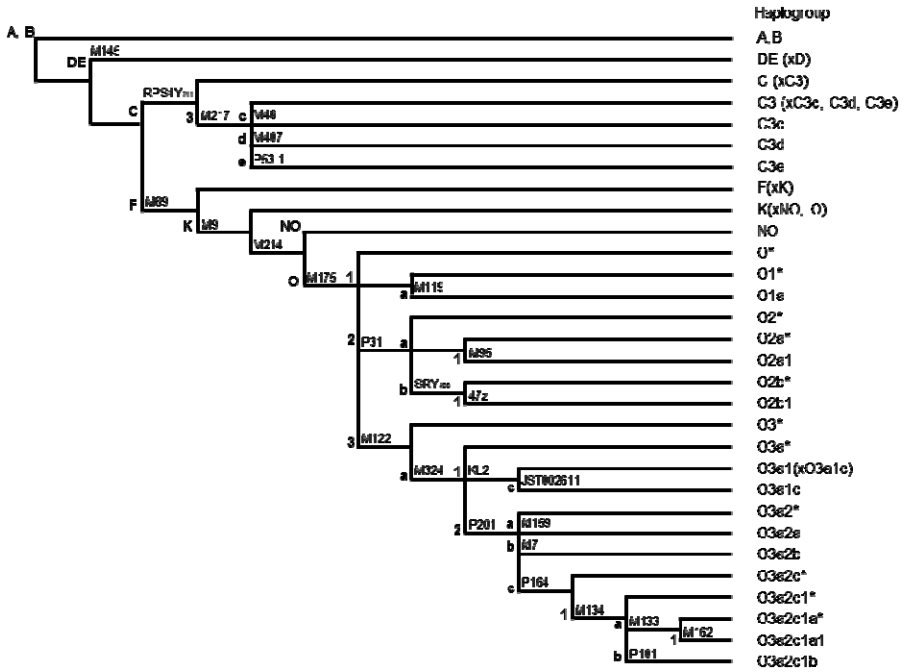


Figure 1. Phylogenetic tree of the 27 Y-chromosomal binary polymorphisms constructed in multiplex SBE reactions. The analyzed Y-SNPs are shown in each branch, and the corresponding haplogroups are shown at the end of each branch according to Karafet et al.[29] and Yan et al.[30]

Primers for PCR amplification and subsequent SBE were designed using Web-based programs Primer 3 (http://frodo.wi.mit.edu/primer3/) and Batchprimer 3 (http://batchprimer3.bioinformatics.ucdavis.edu/index.html), respectively, based on sequences obtained from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) or dbSNP (http://www.ncbi.nlm.nih.gov/snp/). PCR primers were then selected to produce

amplicons smaller than 100 bp and to amplify male-specific fragments in tests using 1 ng of each male and female DNA. SBE primers were designed to provide even peak heights and appropriate peak intervals by adjusting T_m values and adding poly-T and/or poly-A tails to the 5′ ends of primers. PCR and SBE primer sequences are shown in Tables 1 and 2, respectively.

(B) Multiplex PCR amplification and PCR product purification

A total of six multiplex PCR systems were developed. Multiplex PCR amplifications were performed in a final volume of 25 μl that contained 1 ng of template DNA, 2.5 μl of Gold ST*R 10× buffer (Promega), 2.0 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems), and appropriate concentrations of primers (Table 1). Regarding multiplex IV, 1.5 U of AmpliTaq Gold DNA Polymerase were used. Thermal cycling was performed on a Veriti 96-Well Thermal Cycler (Applied Biosystems) under the following conditions: 95°C for 11 min; 33 cycles of 94°C for 20 sec, 60°C for 1 min and 72°C for 30 sec; and a final extension of 72°C for 7 min. For the following SBE reaction, 5.0 μl of the PCR product was purified by incubating at 37°C for 45 min with 1.0 μl of ExoSAP-IT (USB, Cleveland, OH, USA). The enzyme was then inactivated at 80°C for 15 min.

Table 1. Information for the 27 Y-SNP markers, primer sequences and final concentrations for the six multiplex PCR systems

Marker	GenBank/dbSNP accession	PCR primer sequnces (5′→3′)		Conc. (μM)	Amplicon size (bp)
		Forward	Reverse		
Multiplex SBE-I					
M145	rs3848982	gcctccacgactttcctagac	aggttcctcccactccttttt	0.2	93
$RPS4Y_{711}$	rs35284970	cagggcaataaaccttggat	gtggccagcctcttatctctc	1.0	86
M89	rs2032652	agcttcctggattcagctctc	caggatcaccagcaaaggtag	0.2	93
M9	rs3900	ggaccctgaaatacagaactgc	cgtttgaacatgtctaaattaaagaaaa	1.0	85
M214	rs2032674	cactggaaagaaaaagaatgctg	agcctgggagacagt gtgag	0.4	99
M175	rs2032678	acccaaatcaactcaactcca	tgataccttgtttctgttcattctt	0.6	96
Multiplex SBE-II					
M119	AC010889.3	caaaccgcagtgctatgtgt	atgggttattccaattcagca	0.2	94
P31	AC002992.1	tggggaacaggtaggtggta	gtgtgagactccatcgcaaa	0.8	88
M95	AC010889.3	gggatcaaatggagttcctg	gcctacaggttggaaaggcta	0.2	79
SRY_{465}	rs11575897	atcccgcttcggtactctg	tcttgagtgtgtggctttcgt	0.2	83
47z	AC019058.4	tctcctgacctygtgattcg[a]	tcattgacatgggctggact	1.0	87
M122	AC010889.3	cttagttgccttttggaaatgaa	gctttattcagattttcccctga	0.4	88
Multiplex SBE-III					
M324	AC006157.2	tgatttgatctacctgcccttt	aagggaacaaattgatttccag	0.4	70
P201	AC006157.2	tgtgctgtgcaagttgtgtg	tgggtgcagttaagcaatga	0.5	97
M159	AC009977.4	ttcagccttcttctggtacttttta	tcctctggagtcgaaagagtg	0.5	99
M7	AC010084.3	caaagggcatgtaatcattyct[b]	tgatccaattatttccattgtgtt	0.5	100
M134	AC010137.3	atcaaacccagaagggttaaaga	gagatactttgatccccacca	0.4	72
M133	AC010137.3	aaggtggggctttctgaag	gattgtctggttgtggggaa	0.8	98
Multiplex SBE-IV					
KL2	rs17323322	ataccatctgtcctcatccatt	tttaacagaccggcatttgg	0.2	96
JST002611	rs2075181	aggccctgtgcttccagag	atgacccttgcagtgcctt	0.8	93
P201	AC006157.2	tgtgctgtgcaagttgtgtg	tgggtgcagttaagcaatga	0.2	97
P164	rs17316007	gaaccggtgttttctttagca	ccctctttttcctcccattc	0.6	95
M134	AC010137.3	atcaaacccagaagggttaaaga	gagatactttgatccccacca	0.4	
Multiplex SBE-V					
P101	AC010889.3	tctcctaaccttgtgatctgcc	aggaacaccattatctttttcagc	0.2	97
M162	AC009491.3	caggaaaatagatgcctgcaa	cctgacaacagagacagcaca	0.4	86
M133	AC010137.3	aaggtggggctttctgaag	gattgtctggttgtggggaa	0.5	98
Multiplex SBE-VI					
M217	rs2032668	ggagaatgaaaaagttgggtg	aagctgctgtggctttcatc	0.4	82
M48	AC009977.4	tcccttccactcttagcttgac	ctgagggcaactattaaggca	1.0	93
M407	AC006040.3	ctgaaagttggggacagtcat	tggcactaaatcaacttctcctt	0.5	82
P53.1	AC002992.1	caacgaggctgcaggtctta	gaaccaatcccaccctatca	0.6	100

[a] A degenerate forward primer was designed based on sequence alignments of NCBI sequences for the 47z marker (accession number AC019058.4 and NW001842422.1). The degenerate site in the primer sequences is underlined.

[b] A degenerate forward primer was used to amplify samples whether they have a mutation in each primer sequence or not. The degenerate site in the primer sequences is underlined.

Table 2. Primer sequences and final concentrations for the 27 Y-SNP markers in the six multiplex SBE reactions

Marker	Mutation	Orientation[a]	SBE primer sequence (5′→3′)[b]	Primer size (bp)	Conc. (μM)
Multiplex SBE-I					
M145	G/A	Forward	(t)ctagacaccagaaagaaaggc	22	0.10
$RPS4Y_{711}$	C/T	Forward	$(t)_9$agggcaataaaccttggatttc	31	1.00
M89	C/T	Reverse	$(t)_{17}$caactcaggcaaagtgagagat	39	0.16
M9	C/G	Forward	$(t)_{27}$aacggcctaagatggttgaat	48	0.11
M214	T/C	Forward	$(t)_{29}$tggttactttcgttcgtttatttttc	55	1.00
M175	-5bp	Forward	$(t)_{42}$gcacatgccttctcacttctc	63	0.26
Multiplex SBE-II					
M119	A/C	Reverse	ggttattccaattcagcatacaggc	25	1.00
P31	T/C	Forward	$(a)_3$ggttacataaataaggttttttttggttg	33	0.35
M95	C/T	Forward	$(t)_{12}$tgaggataaggaaagactaccatattagtg	42	0.25
SRY_{465}	C/T	Reverse	$(t)_{29}$cctgttgtccagttgcacttc	50	0.10
47z	G/C	Reverse	$(t)_{37}$tgggctggacttggtggctca	58	1.00
M122	T/C	Reverse	$(t)_{45}$ttcagattttcccctgagagc	66	0.40
Multiplex SBE-III					
M324	G/C	Forward	(t)$(a)_2$tgatctacctgccctttcct	23	0.10
P201	T/C	Forward	$(t)_2$(a)gatcttggttaagtcatttgatctcag	30	0.10
M159	A/C	Forward	$(t)_{13}$agttttattattgatgcaagccctaa	39	0.15
M7	C/G	Reverse	$(t)_{14}$(a)ttaaattttgtagttgagttactgttcttctt	47	0.20
M134	-1bp	Forward	$(t)_{34}$agaaaaggcccaggaaagtat	55	0.10
M133	-1bp	Forward	$(a)_{29}$tggggctttctgaagcaaataccagctttaaaaaaaaa	67	1.00
Multiplex SBE-IV					
KL2	T/C	Reverse	$(t)_2$accggcatttgggaactac	21	0.10
JST002611	C/T	Forward	$(t)_{13}$gccctgctagtaggcacca	32	0.18
P201	T/C	Forward	$(t)_{10}$(a)gatcttggttaagtcatttgatctcag	38	0.65
P164	A/G	Forward	$(t)_{23}(a)_2$tagcattttggtcccatctttt	47	0.20
M134	-1bp	Forward	$(t)_{34}$agaaaaggcccaggaaagtat	55	0.12
Multiplex SBE-V					
P101	G/A	Reverse	ggtggctcccgtctgtaatc	20	0.10
M162	C/ C/T	Forward	$(a)_7(t)_1$ggaaaatagatgcctgcaaaa	30	0.80
M133	-1bp	Forward	$(a)_{29}$tggggctttctgaagcaaataccagctttaaaaaaaaa	67	0.60
Multiplex SBE-VI					
M217	A/C	Forward	agaatgaaaaagttgggtgacac	23	0.30
M48	A/G	Forward	agcttgacaattaggattaagaatatgat	29	0.80
M407	A/G	Reverse	$(t)_{14}$gcactaaatcaacttctcctttgg	38	0.80
P53.1	T/C	Reverse	$(t)_{25}$caccctatcactatgcttgtctc	48	0.12

[a] Detection orientation was probed relative to the mutation information reported by Karafet et al. [29]

[b] The tails added to the 5′ ends of SBE primers are in parentheses.

(C) Multiplex SBE reactions and SNP scoring

Six multiplex SBE reactions were carried out with each purified multiplex PCR product, SBE primer mix (Table 2) and a SNaPshot™ Multiplex kit (Applied Biosystems) according to the manufacturer's instructions. Thermal cycling was performed on a Veriti 96-Well Thermal Cycler with 25 cycles at 96°C for 10 sec, 50°C for 5 sec and 60°C for 30 sec. After the SBE reaction, 1.0 U of SAP (USB) was added to the extension product, and the mix was incubated at 37°C for 45 min to remove the unincorporated ddNTPs. SAP was inactivated by incubation at 80°C for 15 min.

The final products were mixed with GeneScan™ 120 LIZ® size standard (Applied Biosystems) and analyzed by capillary electrophoresis with an ABI PRISM 310 Genetic Analyzer (Applied Biosystems) and GeneMapper® ID software 3.2 (Applied Biosystems). SNPs were scored when their peak heights were above the interpretational threshold of 100 relative fluorescent units.

C. Multiplex AS-PCR

(A) Y-SNP selection and primer design for AS-PCR amplification

A set of 19 biallelic Y chromosome markers (M7, M9, M95, M117, M119, M122, M134, M174, M175, M231, M324, P31, P164, P201, JST002611, KL2, $RPS4Y_{711}$, SRY_{465} (M176), and 47z) was selected and the haplogroup tree defined by these markers are shown in Figure 2.

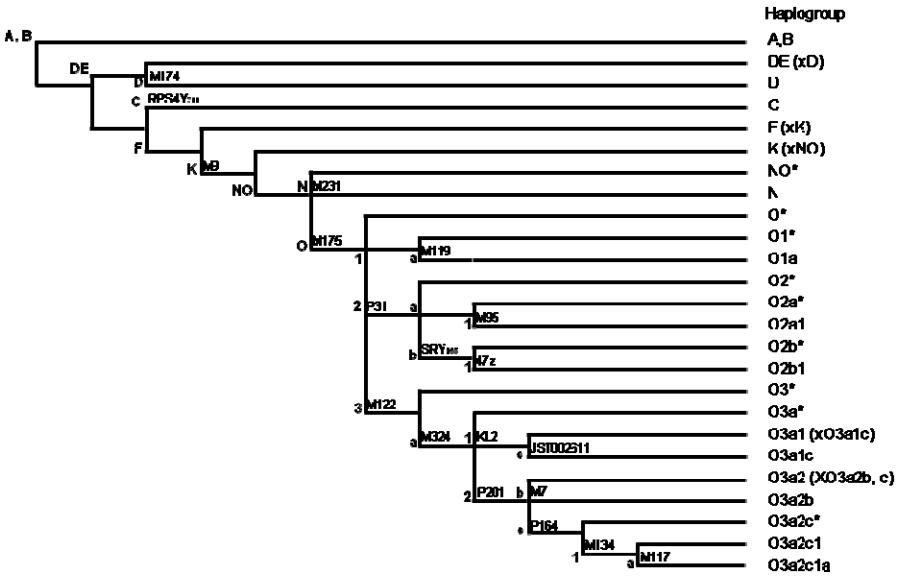


Figure 2. Phylogenetic tree of the 19 Y-chromosomal binary polymorphisms constructed in multiplex AS-PCR assays. The analyzed Y-SNPs are shown in each branch, and the corresponding haplogroups are shown at the end of each branch according to Karafet et al.[29] and Yan et al.[30]

Generic and allele-specific primers were designed using Web-based programs Primer 3 (http://frodo.wi.mit.edu/primer3/) and Batchprimer 3 (http://batchprimer3.bioinformatics.ucdavis.edu/index.html), respectively, based on sequences obtained from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) or dbSNP (http://www.ncbi.nlm.nih.gov/snp/). Allele-specific primers were designed to have a tail at the 5′ end, thereby allowing different alleles to produce amplicons of different sizes. Generic and allele-specific primer sequences are shown in Table 3.

Table 3. Information for the 19 Y-SNP markers, allele specific primer sets and final concentrations for multiplex allele-specific PCR system

Marker	GenBank/dbSNP accession	Mutation		PCR primer sequnces (5′→3′)[a]	Conc. (μM)	Amplicon size (bp)
Multiplex AS-I						
M174	rs2032602	T	F	(aa)gaataccttctggagtgccc**t**	0.30	81
		C	F	aataccttctggagtgccc**c**	0.30	78
		-	R	FAM-aggaaaagtgtgcaataaacactataa	0.30	-
$RPS4Y_{711}$	rs35284970	C	F	ggcaataaaccttggatttc**c**	0.20	102
		T	F	(at)gggcaataaaccttggatttc**t**	0.10	105
		-	R	FAM-ttagccactgctctgtttggt	0.20	-
M9	rs3900	C	F	(tt)aacggcctaagatggttgaat**c**	0.85	130
		G	F	(g)cggcctaagatggttgaat**g**	0.50	127
		-	R	FAM-tttgagtattgaaatgcataatgaag	0.85	-
M231	rs9341278	-	F	FAM-cctctcacaaaatggattctgatt	0.60	-
		G	R	(aaa)gacgatctttcccccaatt**c**	0.20	156
		A	R	gacgatctttcccccaatt**t**	0.60	153
M175	rs2032678	-	F	aactctctgaatcaggcacatg	0.55	80
		-5bp	R	VIC-ccttttttttctactgataccttgttt	0.55	75
M119	AC010889.3	-	F	VIC-tttcagagtttatgctccaaacc	0.30	-
		A	R	(at)ttccaattcagcatacaggc**t**	0.30	105
		C	R	(g)ccaattcagcatacaggc**g**	0.10	102
M122	AC010889.3	-	F	VIC-caaatggtatgcaactcagctatatt	0.55	-
		T	R	(aa)cagattttcccctgagagc**a**	0.20	131
		C	R	agattttcccctgagagc**g**	0.55	128
JST002611	rs2075181	C	F	(att)agccctgctagtaggcac<u>a</u>a**c**	0.50	159
		T	F	agccctgctagtaggca<u>a</u>ca**t**	0.30	156
		-	R	VIC-tactcacaggggcattgca	0.50	-
P31	AC002992.1	T	F	(aa)tggttacataaataaggttttc<u>t</u>ttttggttg**t**	1.00	81
		C	F	ggttacataaataaggttttc<u>t</u>ttggttg**c**	1.00	78
			R	NED-agcctgggcaacagtgtga	1.00	-
M95	AC010889.3	-	F	NED-cagtcagctgccagcaatagt	1.00	-
		C	R	(taa)ggaaaggctaagccatc<u>g</u>a**g**	1.00	105
		T	R	ggaaaggctaagccatgca**a**	0.60	102
SRY_{465}	rs11575897	-	F	NED-gcacagagagaaatacccgaat	0.95	-
		C	R	(tt)cctgttgtccagttgcact<u>a</u>c**g**	0.50	132
		T	R	ctgttgtccagttgcac<u>g</u>tc**a**	0.95	129
47z	AC019058.4	G	F	(a)aaagtgctgggattaca<u>t</u>gc**g**	0.50	159
		C	F	agtgctgggattacaggc**c**	0.35	156
		-	R	NED-gaaataatacaaagaaatcctgcct	0.50	-
P201	AC006157.2	T	F	(t)tcttggttaagtcatttgatctgag**t**	0.75	82
		C	F	ttggttaagtcatttgatctc<u>t</u>g**c**	0.45	79
		-	R	PET-agtgctgggtgcagttaagc	0.75	-
M7	AC010084.3	-	F	PET-tagattc<u>g</u>tgagggcagca	0.40	-
		C	R	aattttgtagttgagttactgttct<u>g</u>ctt**g**	0.40	103
		G	R	(a)taaattttgtagttgagttactgttc<u>g</u>tctt**c**	0.40	106
M134	AC010137.3	-	F	PET-aaaaatttccccacaaccaga	0.20	128
		-1bp	R	agaattctcatttaccactgtggag	0.20	127
M117	AC010889.3	-	F	PET-aggctattaatcaaaaataccagca	1.00	158
		-4bp	R	ttaaaattgacag<u>w</u>tatcagtttgaaatta[b]	2.00	154

Table 3 Cont.

Marker	GenBank/dbSNP accession	Mutation		PCR primer sequnces (5′→3′)[a]	Conc. (μM)	Amplicon size (bp)
Multiplex AS-II						
M324	AC006157.2	-	F	FAM-aggaacatagcaagacccaaaa	0.30	-
		G	R	(g)acatgggctgcaacgaga**c**	0.30	103
		C	R	(ttt)tacatgggctgcaacagga**g**	0.30	106
KL2	rs17323322	-	F	VIC-cctcatccatttaaaggccaa	0.15	-
		T	R	(tt)accggcatttgggaactac**a**	0.10	79
		C	R	ccggcatttgggaactac**g**	0.15	76
JST002611	rs2075181	C	F	(att)agccctgctagtaggcac<u>aa</u>**c**	0.35	159
		T	F	agccctgctagtaggca<u>ac</u>**t**	0.20	156
		-	R	VIC-tactcacaggggcattgca	0.35	-
P201	AC006157.2	T	F	(t)tcttggttaagtcatttgatct<u>g</u>ag**t**	0.90	82
		C	F	ttggttaagtcatttgatctc<u>t</u>g**c**	0.60	79
		-	R	PET-agtgctgggtgcagttaagc	0.90	-
P164	rs17316007	A	F	(aa)agcattttggtcccatct<u>g</u>tt**a**	0.80	106
		G	F	gcattttggtcccatcttt**g**	0.40	103
		-	R	PET-agctttttctcaattattcatgct	0.80	-
M134	AC010137.3	-	F	PET-aaaaatttccccacaaccaga	0.20	128
		-1bp	R	agaattctcatttaccactgtggag	0.20	127

[a] Two alleles of a single-nucleotide polymorphism (SNP) at the 3′ end of allele-specific primers are indicated in bold; additional mismatches from the 3′ end to enhance the specificity in the allele-specific PCR are underlined; the tails inserted at the 5′ end of one of the two allele-specific primers to produce amplicons of different sizes are in parentheses.

[b] A degenerate reverse primer was designed to amplify samples whether they have a SNP (M116.1) in the primer sequence or not. The SNP site in the primer sequences is underlined.

(B) Multiplex PCR amplification and SNP scoring

Two multiplex allele-specific PCR assays were developed. Multiplex PCR reactions for all markers except the 47z marker were performed in a final volume of 10 μl containing 1 ng of template DNA, 1.0 μl of Gold ST*R 10× buffer (Promega), and appropriate concentrations of primers (Table 3) with 2.5 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems) for multiplex AS-I and 1.5 U for multiplex AS-II. Because of the difficulty in simultaneous amplification of the 47z marker with other markers in the multiplex AS-I,

monoplex allele-specific PCR was performed for the 47z marker with the same PCR conditions as above except 0.5 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems) was used. Thermal cycling was done on a Veriti 96-Well Thermal Cycler (Applied Biosystems) with the following conditions: 95°C for 11 min; 30 cycles of 94°C for 20 sec, 59°C for 1 min, and 72°C for 30 sec; and a final extension of 60°C for 45 min. The PCR product of the 47z marker was mixed with an equal amount of product from the multiplex AS-I before being separated by capillary electrophoresis using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems) after mixing with GeneScan™ 500 LIZ® size standard (Applied Biosystems). Automated allele calling was performed using GeneMapper® ID software 3.2 (Applied Biosystems). An allelic ladder containing both ancestral and derived alleles for all markers was used to perform allelic designation in the multiplex allele-specific PCR assays.

D. PCR amplification with serially diluted DNA and degraded DNA

Five replicates of serially diluted 9948 male DNA or 2800M Control DNA were amplified independently using multiplexes SBE-I, II, III, IV, V, VI, AS-I and AS-II under the same PCR conditions described above. Additional five replicates of serially diluted 9948 male DNA or 2800M Control DNA for multiplexes SBE-I, II, III, IV, V and VI were then amplified independently under the same PCR conditions except for the number of PCR cycles: 33 cycles for 1 ng, 500 pg, and 250 pg of template DNA; 35 cycles for 125 and 62 pg; and

37 cycles for 31 and 15 pg. The genotype results for each diluted DNA sample were compared to known genotypes that were determined with 1 ng of the standard DNA. To test the effect of input DNA amount on SNP scoring, the peak height of each amplicon was obtained with GeneMapper software, and average and standard deviation values were calculated at each concentration.

One microliter of artificially degraded DNA was independently amplified twice using multiplexes SBE-I, II, III, IV, V, VI, AS-I and AS-II under the same PCR conditions as multiplex PCR amplification of 1 ng DNA. The genotyping results were then confirmed to be identical to those of 1 ng non-degraded human genomic DNA. Ten DNA samples that were extracted from 55-year-old skeletal remains[35] were also amplified with at least 60 pg of template DNA using multiplex SBE-I, II, III, IV, V and VI under the same PCR conditions except for using 35 PCR cycles. Then, following the low copy number DNA interpretation rule (i.e. replicate analyses with duplicate results prior to reporting alleles),[39, 40] SNPs were scored only when the SNP was observed in common from two replicate reactions.

E. Monoplex PCR followed by SBE reaction

Monoplex PCR followed by SBE reaction was also used to analyze M110, M133, M159, M207, M242, M267, P203 and PK4 markers. The monoplex PCR reactions were performed in a final volume of 25 μl which contained 1 ng of template DNA, 2.5 μl of Gold ST*R 10× buffer (Promega), 2.0 U of AmpliTaq

Gold® DNA polymerase (Applied Biosystems), and appropriate concentrations of primers (Table 4). Thermal cycling was done with the following conditions: 95°C for 11 min; 33 cycles of 94°C for 20 sec, 60°C for 1 min, and 72°C for 30 sec; and a final extension of 72°C for 7 min. Before the SBE reaction, 5.0 μl of the PCR product was purified by incubating at 37°C for 45 min with 1.0 μl of ExoSAP-IT (USB, Cleveland, OH, USA). The enzyme was inactivated at 80°C for 15 min. The SBE reaction was carried out with a SNaPshot™ Multiplex kit (Applied Biosystems), the purified PCR product, and SBE primer mix (Table 5) according to the manufacturer's instructions. Thermal cycling conditions were as follows: 25 cycles at 96°C for 10 sec, 50°C for 5 sec, and 60°C for 30 sec. After the SBE reaction, 1 U of SAP (USB) was added to the extension product, and the mix was incubated at 37°C for 45 min to remove the unincorporated ddNTPs. SAP was inactivated by incubating at 80°C for 15 min. The purified SBE products were separated by capillary electrophoresis using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems) after mixing with GeneScan™ 120 LIZ® size standard (Applied Biosystems). Automated allele calling was made using GeneMapper® ID software 3.2 (Applied Biosystems).

Table 4. Information for the 8 Y-SNP markers, primer sequences and final concentrations for the monoplex PCR system

Marker	GenBank/dbSNP accession	PCR primer sequnces (5′→3′)		Conc. (μM)	Amplicon size (bp)
		Forward	Reverse		
M110	AC010889.3	gacgtttaggaggcaggatg	aaaatccaacgacaaatgtgc	0.6	99
M133	AC010137.3	aaggtggggctttctgaag	gattgtctggttgtggggaa	0.6	98
M159	AC009977.4	ttcagccttcttctggtactt ttta	tcctctggagtcgaaagagtg	0.6	99
M207	rs2032658	gggcaaatgtaagtcaagcaa	gccaattaggtcacttcaacc	0.6	92
M242	rs8179021	gcaaaaaggtgaccaaggtg	gaacaactctgaagcggtgg	0.6	114
M267	rs9341313	ggattcgatggaagcatttt	ggattgcaggcatcagcta	0.6	80
P203	Rs13447354	tggttggctattgagttagcat	ttgtgctgtatttttcatttgtttt	0.6	117
PK4	AC007244.2	agaaattccaaaatcggcaa	ccacaaagcactccaacaaa	0.6	98

Table 5. Primer sequences and final concentrations for the 8 Y-SNP markers in the SBE reaction

Marker	Mutation	Orientation[a]	SBE primer sequence (5′→3′)[b]	Primer Size (bp)	Conc. (μM)
M110	T/C	Forward	caggatgccggtacaatgtatt	22	0.20
M133	-1bp	Forward	$(a)_{29}$tggggctttctgaagcaaataccagctttaaaaaaaaa	67	0.20
M159	A/C	Forward	$(t)_{13}$agttttattattgatgcaagccctaa	39	0.20
M207	A/G	Forward	$(t)_{19}$gcaaatgtaagtcaagcaagaaattta	46	0.20
M242	C/T	Reverse	$(t)_{19}$acacgttaagaccaatgccaa	40	0.20
M267	T/G	Forward	tcgatggaagcatttttgtaaata	24	0.20
P203	G/A	Forward	gttggctattgagttagcataatca	25	0.20
PK4	A/T	Forward	aatcggcaatggtttttgag	20	0.20

[a] Detection orientation was probed relative to the mutation information reported by Karafet et al.[29]

[b] The tails added to the 5′ ends of SBE primers are in parentheses.

2. Y-STR analysis

A. DNA samples

Blood or buccal samples from 706 unrelated Korean males, including 355 males already typed for 22 Y-STRs[34] and 301 males typed for 19 Y-STRs,[33] were analyzed. Genomic DNA was extracted using a QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. 9948

Male DNA (Promega, Madison, WI, USA) was used to calibrate allelic ladders.

B. Multiplex PCR and genotyping for 22-Y-STRs

A total of three multiplex PCR sets were constructed for 22 Y-STRs (Figure 3). Multiplex STR-I consisted of the minimal haplotype STRs (DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 and DYS385). Multiplex STR-II consisted of DYS385, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635 (GATA C4) and GATA H4.1. Multiplex STR-I and II implemented 17 Y-STRs from the AmpFlSTR® Yfiler™ kit (Applied Biosystems). Multiplex STR-III consisted of DYS385, DYS388, DYS446, DYS447, DYS449, and DYS464. In three multiplexes, DYS385 was commonly used to check for sample switching. PCR amplifications were carried out in a final volume of 10 μl containing 1.0 ng template DNA, 1.6 μl of Gold ST*R 10X buffer (Promega), 2.0 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems) and the appropriate concentration of primers (Table 6). Thermal cycling was conducted on the Veriti 96-Well Thermal Cycler (Applied Biosystems) or PTC-200 DNA engine (MJ Research, Waltham, MA, USA) under the following conditions: 95°C for 11 min; 30 cycles of 94°C for 1 min, 55°C (multiplex STR-I) or 59°C (multiplexes STR-II and III) for 1 min, 72°C for 1 min, and a final extension at 60°C for 45 min.

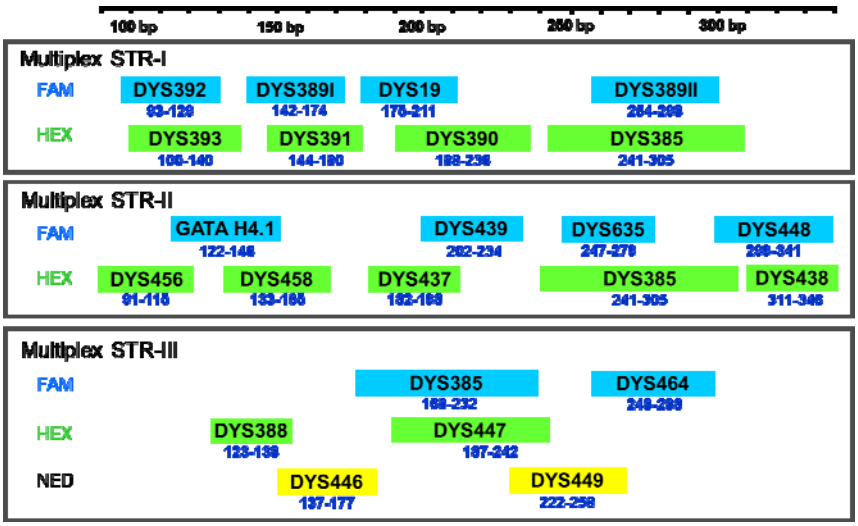


Figure 3. Schematic of three multiplex PCR sets for 22 Y-STR loci.

The PCR products were mixed with GeneScan-400 HD (ROX) size standard (Applied Biosystems) and separated by capillary electrophoresis using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The typing of PCR products at each STR locus was performed by comparing them to an allelic ladder, which was constructed after confirming the sequences. Allele nomenclature followed the recommendations of the International Society of Forensic Genetics (ISFG) Commission[41], and allele designation was carried out using GeneMapper® ID software 3.2 (Applied Biosystems).

Table 6. Primer sequences and concentrations of three multiplex PCR sets for 22 Y-STRs

PCR Set	Locus	Dye	Sequences (5′ to 3′)[a]	Conc.(μM)
Multiplex STR-I	DYS19	FAM	ctactgagtttctgttatagt	0.40
			atggccatgtagtgaggaca	0.40
	DYS385a/b	HEX	agcatgggtgacagagcta	0.60
			ccaattacatagtcctcctttc	0.60
	DYS389I/II	FAM	ccaactctcatctgtattatct	0.20
			ttatccctgagtagtagaagaat	0.20
	DYS390	HEX	tatattttacacatttttgggcc	0.10
			tgacagtaaaatgaacacattgc	0.10
	DYS391	HEX	ctattcattcaatcatacacccatat	0.06
			acatagccaaatatctcctggg	0.06
	DYS392	FAM	aaaagccaagaaggaaaacaaa	0.06
			aaacctaccaatcccattcctt	0.06
	DYS393	HEX	gtggtcttctacttgtgtcaatac	0.05
			aactcaagtccaaaaaatgagg	0.05
Multiplex STR-II	DYS385a/b	HEX	agcatgggtgacagagcta	0.30
			gccaattacatagtcctcctttc	0.30
	DYS437	HEX	gactatgggcgtgagtgcat	0.20
			gagaccctgtcattcacagatga	0.20
	DYS438	HEX	ccaaaattagtggggaatagttg	0.25
			gatcacccagggtctggagtt	0.25
	DYS439	FAM	tcgagttgttatggttttaggtct	0.18
			gtggcttggaattcttttaccc	0.18
	DYS448	FAM	tgggagaggcaaggatccaa	0.70
			gtcatatttctggccggtctgg	0.70
	DYS456	HEX	ggaccttgtgataatgtaagatag	0.10
			gtagagggacagaactaatggaa	0.10
	DYS458	HEX	gcaacaggaatgaaactccaat	0.10
			gttctggcattacaagcatgag	0.10
	DYS635	FAM	agtgtctcacttcaagcaccaagcac	0.45
			gcagcaaaattcacagttggaaaaatgt	0.45
	GATA H4.1	FAM	atgctgaggagaatttccaa	0.22
			gctattcatccatctaatctatccatt	0.22
Multiplex STR-III	DYS385a/b		gaaggaaggaaggaagggaaa	0.05
		FAM	taagggctgctgaccagatt	0.05
	DYS388	HEX	gtgagttagccgtttagcga	0.16
			gcagatcgcaaccactgcg	0.16
	DYS446		gtattttcagtcttgtcctgtc	0.30
		NED	gagcttgtaccactgcactca	0.30
	DYS447	HEX	ggtcacagcatggcttggtt	0.08
			gggcttgctttgcgttatctct	0.08
	DYS449	NED	cttgctctttttctttctctctt	0.42
			gcactctaggttggacaacaa	0.42
	DYS464	FAM	ttacgagctttgggctatg	0.68
			gcctgggtaacagagagactctt	0.68

[a] The 5′ nucleotide g of unlabeled primer was added to promote adenylation.

C. PCR amplification and sequence analysis for DYS385

Locus-discrimination primers for DYS385-Fa (AC007379) and DYS385-Fb (AC022486) were designed using the Primer 3 software (www.genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). These primers were used for the analysis of samples with allele designation discrepancies between the different primer pairs of multiplex STR-I/II and multiplex STR-III. To compare these primer pairs with commercial systems, I used the Powerplex® Y (Promega) and the AmpFlSTR® Yfiler™ (Applied Biosystems) kit according to the manufacturer's recommendations.

Sequence analyses were performed on the flanking region of samples with discordant allele designation. Each PCR products was cloned using pGEM®-T Easy Vector System I (Promega) following the manufacturer's recommendation, and then sequenced on an ABI PRISM 310 Genetic Analyzer using a BigDye Terminator Cycle Sequencing v2.0 Ready Reaction kit (Applied Biosystems).

D. Multiplex PCR assay to characterize the null allele at DYS448

To detect rearrangements at the azoospermia factor c (*AZFc*) region, I amplified five *AZFc*-specific sequence tagged site (STS) markers, sY1161, sY1191, sY1201, sY1206 and sY1291, and a control gene pair ZFX/ZFY in samples with a null allele at DYS448. Primers for the five markers and the ZFX/ZFY gene were the same as used by Lin et al.[42] Multiplex PCR reactions

were performed in a total volume of 10 μl containing 10 ng DNA, 2.0 U AmpliTaq Gold DNA polymerase (Applied Biosystems), 1X PCR buffer I (Applied Biosystems), 200 μM each dNTP and appropriate concentrations of primers. Primer concentrations were adjusted empirically to balance PCR yields. Thermal cycling conditions were as follows: 95℃ for 11min; 35 cycles of 94℃ for 30 s, 61℃ for 45 s, 72℃ for 1 min; and a final extension of 72℃ for 10 min. The amplification reaction was analyzed by agarose gel electrophoresis stained with ethidium bromide to observe amplified STS markers.

3. Statistical analysis

The relative frequencies of allele/haplotype occurrences, standard diversity parameters including gene/haplotype diversity and analysis of molecular variance (AMOVA) were calculated with the software package Arlequin 3.5.1.3.[43] Discrimination capacity was estimated as the number of haplotype divided by the number of individuals typed. The haplogroup frequencies were determined by direct counting. AMOVA was performed for STR allele frequencies among haplogroups to test simple hierarchical partitioning of haplotypes in the haplogroup. A median-joining (MJ) network was constructed for the Y-STR haplotypes within specific haplogroups by the program NETWORK 4.6.0.0 (http://www.fluxus-engineering.com). Y-STR weighting was applied in accordance with Gomes et al.[44]

III. RESULTS

1. Development and assessment of multiplex PCR systems for haplogroup typing

A. Multiplex PCR followed by SBE reaction

The selected 22 Y-SNP markers were amplified with four multiplex PCR systems to explore East Asian Y chromosome haplogroups. Multiplex SBE-I was composed of the six Y-SNPs, M145, RPS4Y$_{711}$, M89, M9, M214, and M175, which distinguish all world-wide major clades except for the African lineages A and B. Among the six major haplogroups in multiplex SBE-I, clade O is the most common haplogroup in East Asians, and accordingly, a more detailed analysis of the haplogroup O lineage is necessary to differentiate East Asian Y haplogroups. Therefore, multiplex SBE-II (M119, P31, M95, SRY$_{465}$, 47z, and M122), III (M324, P201, M159, M7, M134, and M133), IV (KL2, JST002611, P201, P164 and M134) and V (M133, M162 and P101) were constructed to subdivide haplogroup O into subhaplogroups O1, O2 and O3 and their internal derivatives (Figure 4). Finally, multiplex SBE-VI (M48, M217, M407, and P53.1) was developed to further define subhaplogroups C, especially haplogroup C3 and its derivatives which are widespread in East Asia .[38]

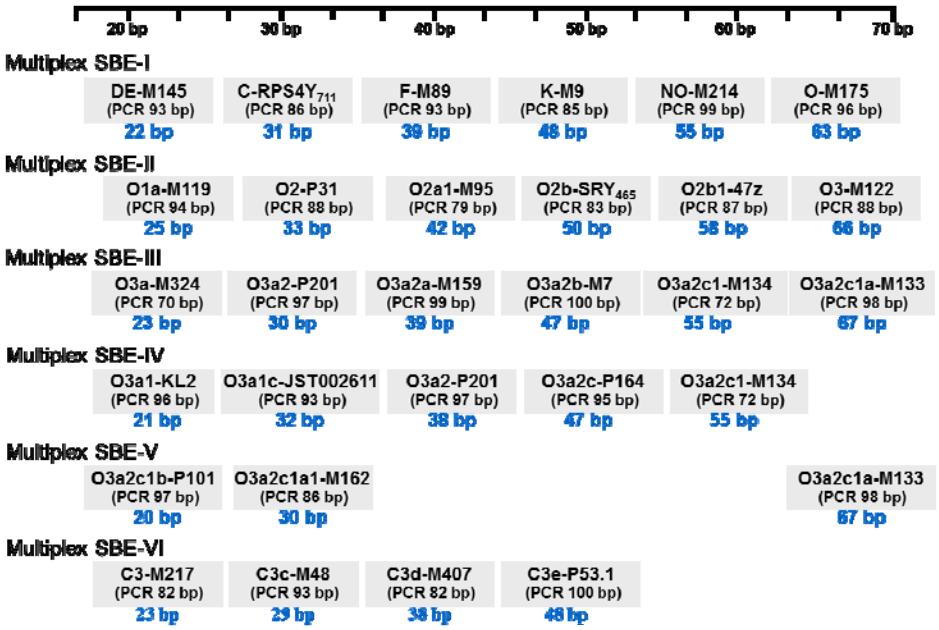


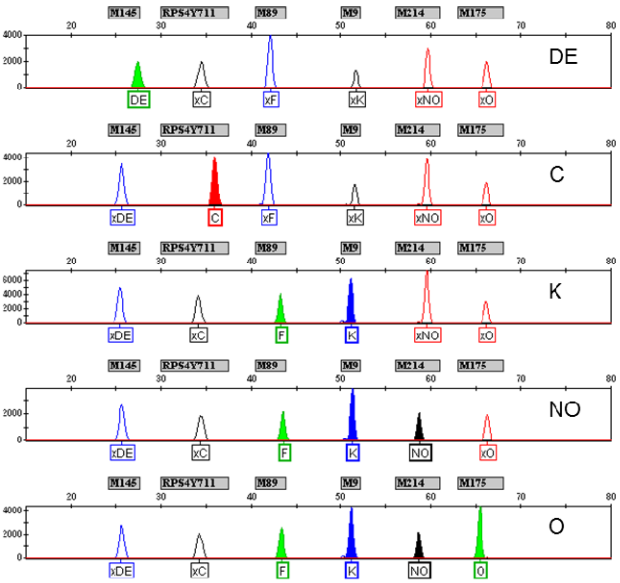
Figure 4. Schematic of six multiplex SBE reactions for genotyping the 27 Y-SNPs.

All PCR primers were designed to produce small amplicons between 70 and 100 bp in length to facilitate the analysis of degraded forensic and ancient samples (Table 1 and Figure 4).[45, 46] For amplification of the 47z marker (DXYS5) that is located in a Y-chromosomal region with high sequence similarity to the X chromosome (DXYS5X), the 3′ end of the reverse PCR primer was designed to be located at a nucleotide position that shows a mismatch between DXYS5Y and DXYS5X sequences, thereby allowing specific amplification of the target sequence on the Y chromosome only (DXYS5Y). In addition, a degenerate forward primer was used to allow the amplification of alternate assembly (based on HuRef, sequence accession number NW001842422.1) as well as the Genbank reference sequence (sequence

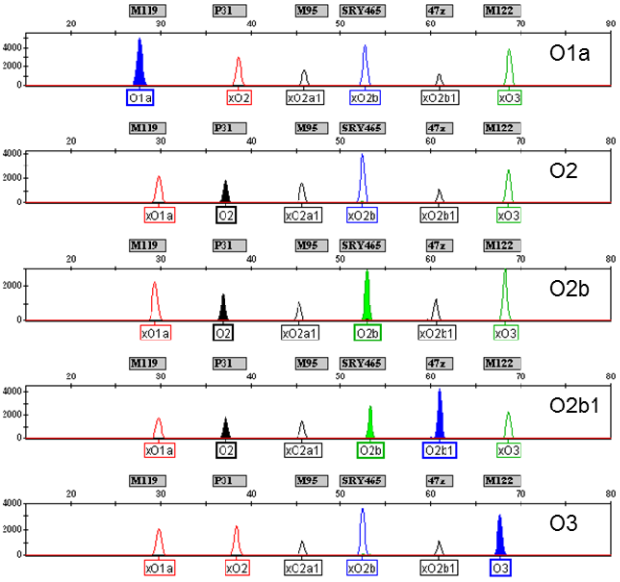
accession number AC019058.4). Male-specific amplification of the 47z marker on the Y chromosome (DXYS5Y) was confirmed in several male and female DNA samples.

SBE primers were designed to produce extension products ranging in size from 22 to 67 bases. Each extension product had a size difference of more than 7 bases from all others, so that the extension products could be easily discriminated during electrophoresis. For this reason, the 5′ ends of certain primers were tailed with poly-T nucleotides. If the complementary nucleotide for the tail was A, nucleotide A was used instead of nucleotide T (Table 1). In case of the M133 marker primer with a 3′ end AAAAAA sequence, a 5′ poly-A tail was used instead of a poly-T tail to avoid self-complementary binding of the primer (Table 1). Representative electropherograms of the six developed multiplexes are shown in Figure 5.

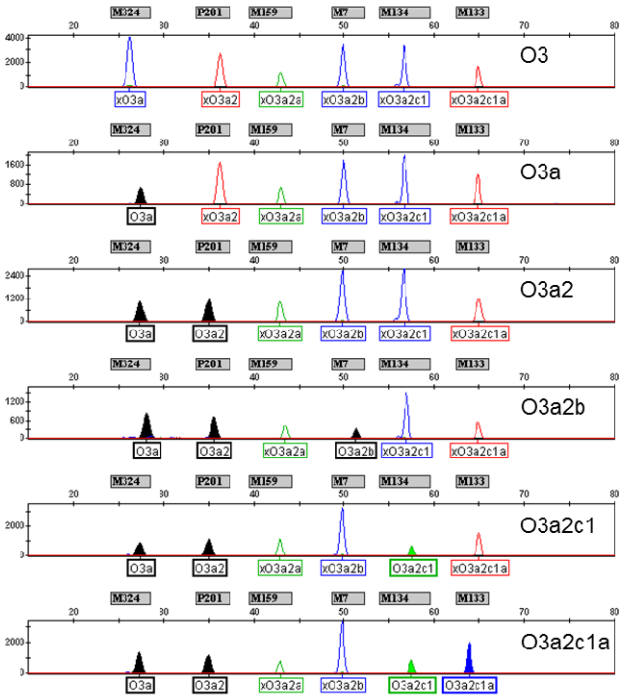
a


M145
RPS4Y711
M89
M9
M214
M175
DE
xC
xF
xK
xNO
xO
C
xDE
K
F
NO
O

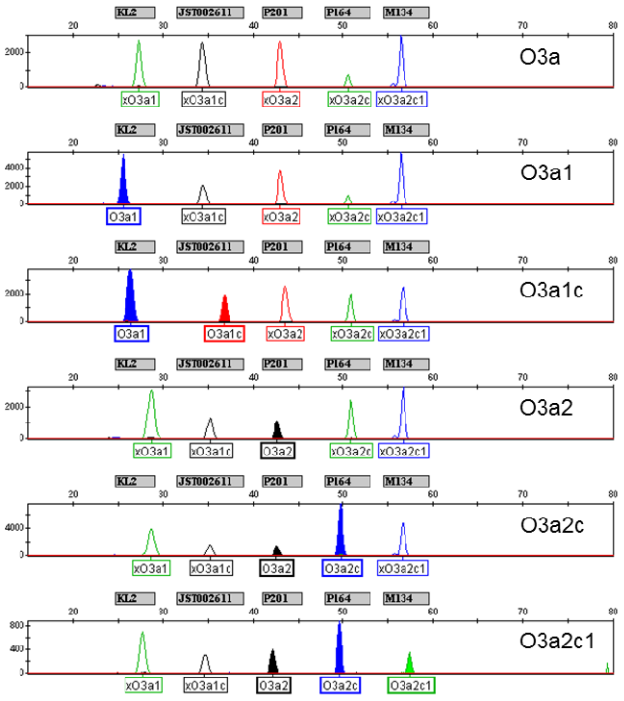
b


M119
P31
M95
SRY465
47z
M122
O1a
xO2
xO2a1
xO2b
xO2b1
xO3
O2
xO1a
xC2a1
O2b
xO2t1
O2b1
O3

c


M324
P201
M159
M7
M134
M133
O3
xO3a
xO3a2
xO3a2a
xO3a2b
xO3a2c1
xO3a2c1a
O3a
O3a2
O3a2b
O3a2c1
O3a2c1a

d


KL2
JST002611
P201
P164
M134
O3a
xO3a1
xO3a1c
xO3a2
xO3a2c
xO3a2c1
O3a1
O3a1c
O3a2
O3a2c
O3a2c1

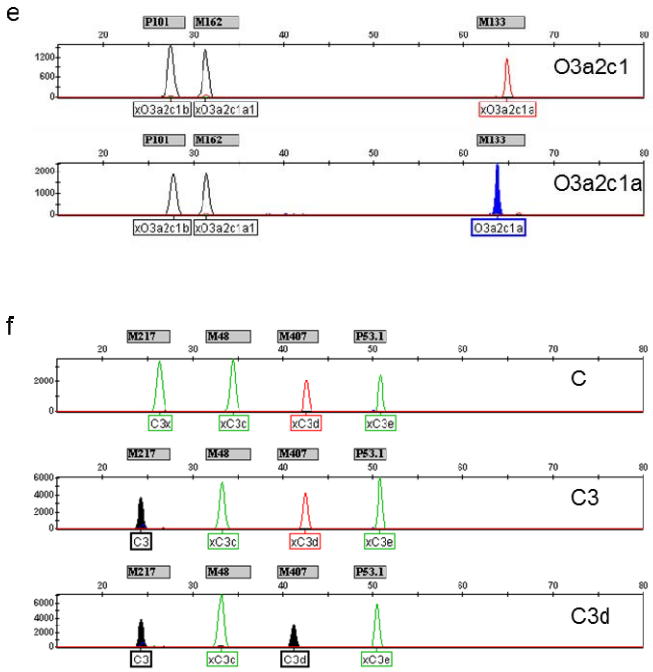


Figure 5. Representative electropherograms of multiplex SBE-I (a), II (b), III (c), IV (d), V (e) and VI (f) obtained from male donors belonging to each haplogroup observed in this study. Each labeled peak represents a SNP and the right-hand side of the each panel shows the relevant haplogroup.

B. Multiplex allele-specific PCR assays

A set of 19 Y-SNPs (M7, M9, M95, M117, M119, M134, M174, M175, M122, M231, M324, P31, P164, P201, JST002611, KL2, $RPS4Y_{711}$, SRY_{465} and 47z) was selected to construct two multiplex allele-specific PCR assays for correspondence to the multiplex SBE reactions. (Figure 6) Multiplex AS-I was consisted of 16 Y-SNPs, M7, M9, M95, M117, M119, M134, M174, M175, M122, M231, P31, P201, JST002611, $RPS4Y_{711}$, SRY_{465} and 47z to determine the comprehensive haplogroups observed frequently in East Asian according to

the tree by Karafet et al.[29] Multiplex AS-II was comprised of M134, M324, P164, P201, JST002611 and KL2 to determine the revised haplogroups following the tree of Yan et al.[30] Three markers (M134, P201, and JST002611) are common to multiplexes AS-I and AS-II, which enabled to check for sample switch and to confirm the typing results. The multiplex allele-specific PCR assays were optimized for simultaneous detection of 22 Y-SNPs (16 in multiplex AS-I and 6 in multiplex AS-II) by fragment analysis on an automatic DNA sequencer like general forensic STR typing method. Representative electropherograms of the two developed multiplex AS-PCR assays are shown in Figure 7.

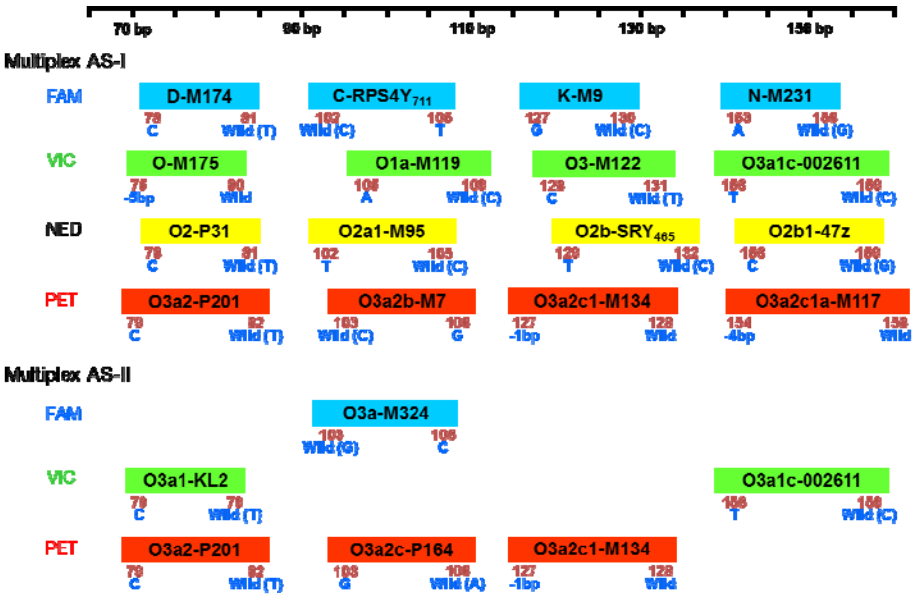
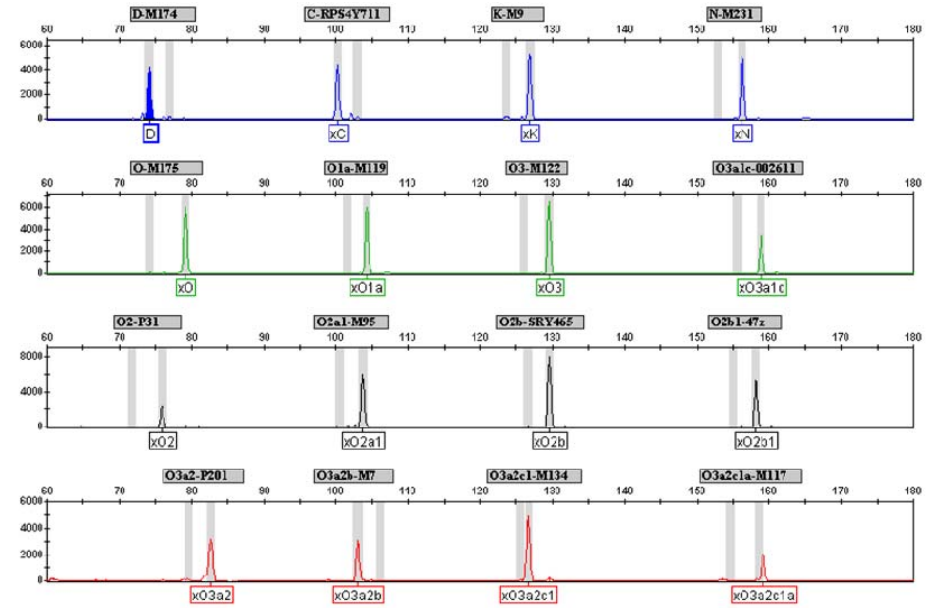


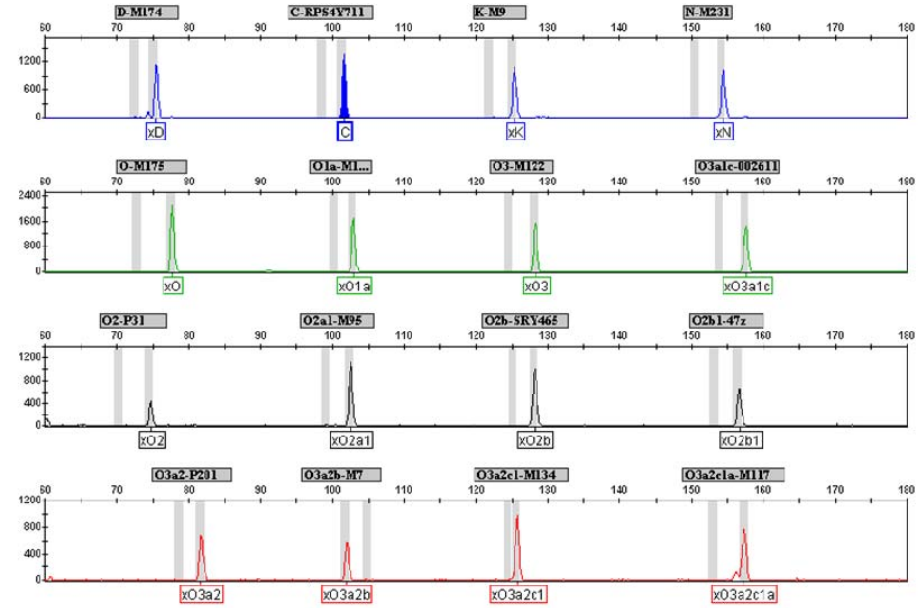
Figure 6. Schematic of two multiplex AS-PCR assays for genotyping the 19 Y-SNPs.

a



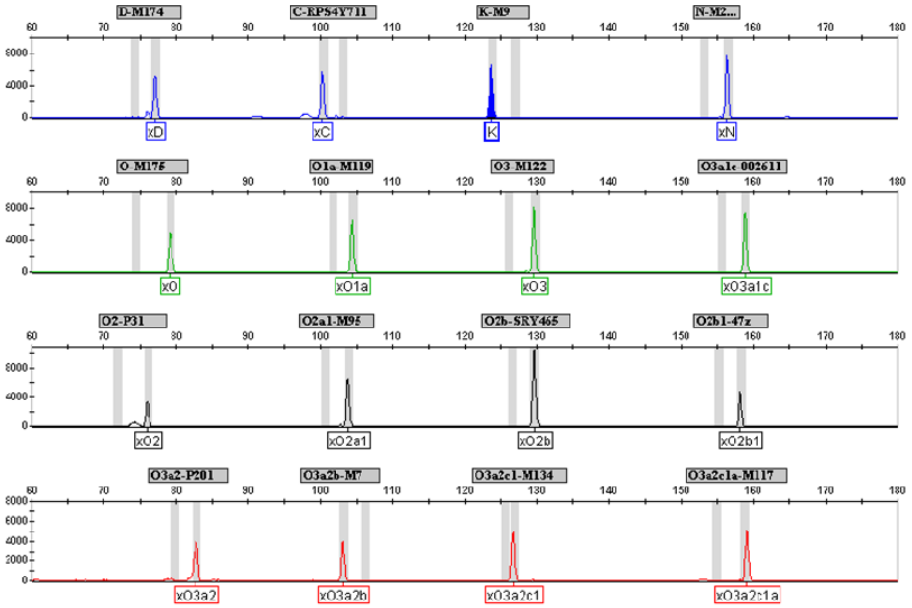
Haplogroup D

b



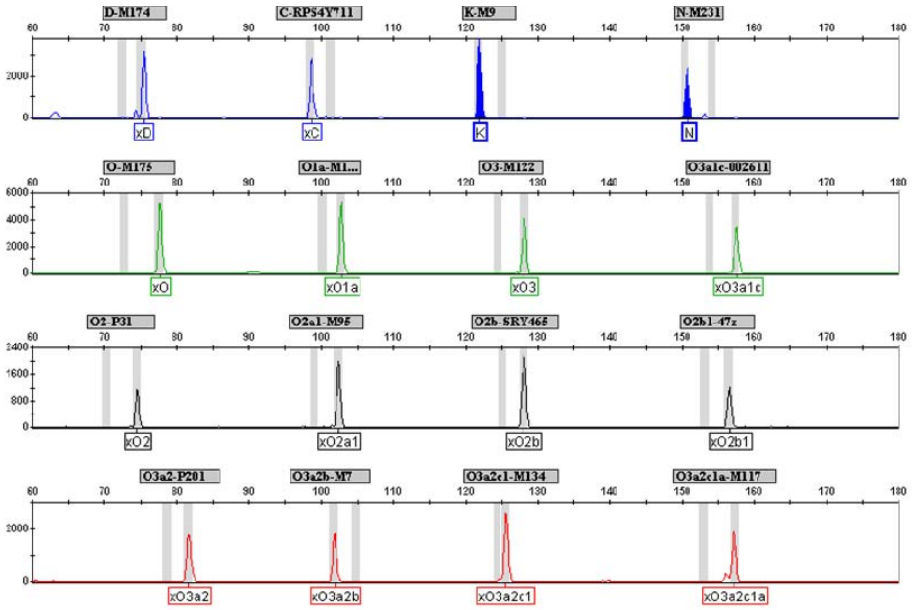
Haplogroup C

c



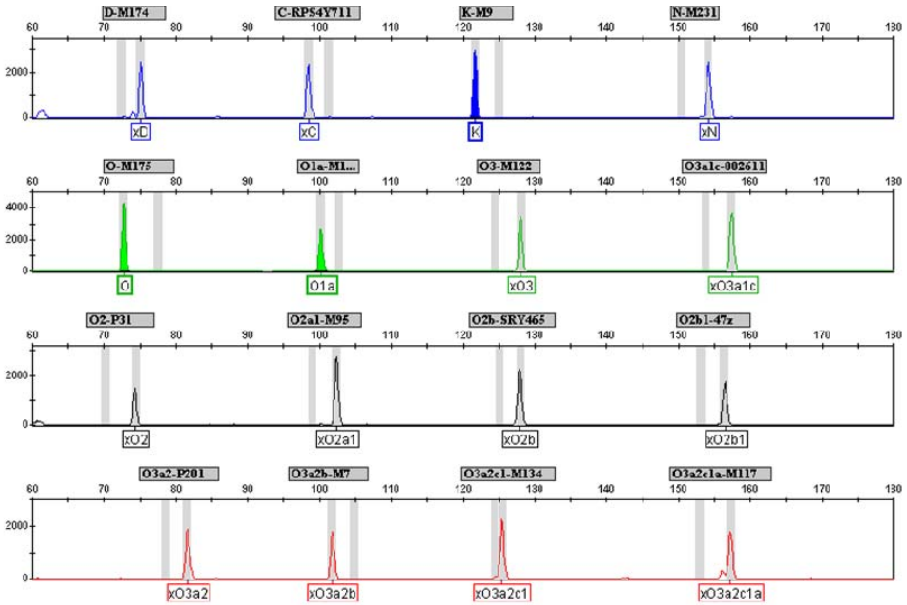
Haplogroup K

d



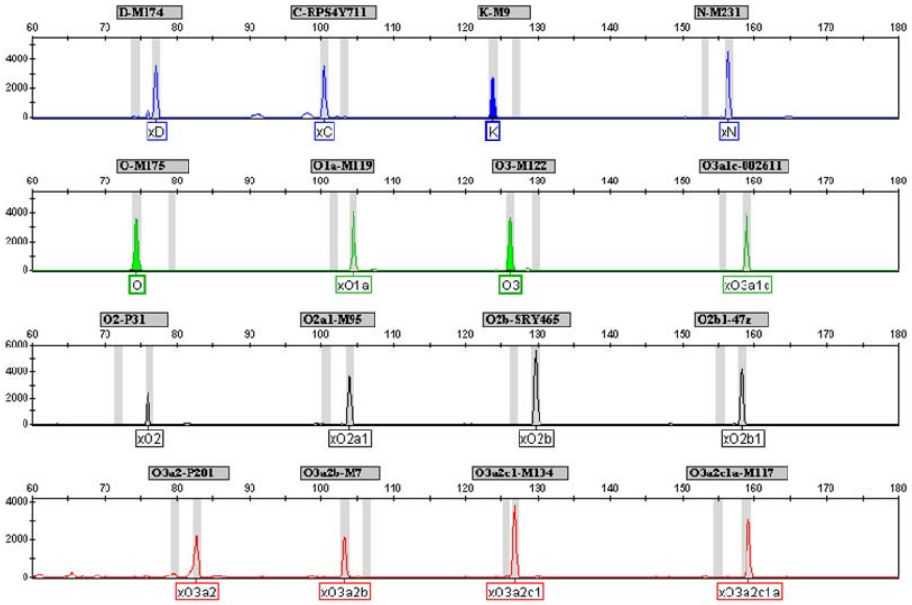
Haplogroup N

e



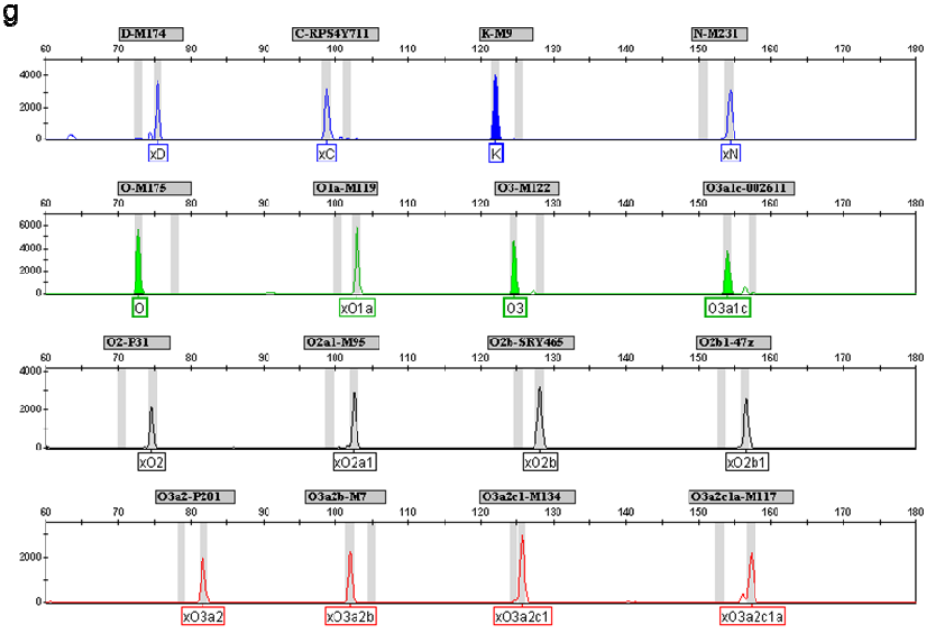
Haplogroup O1a

f



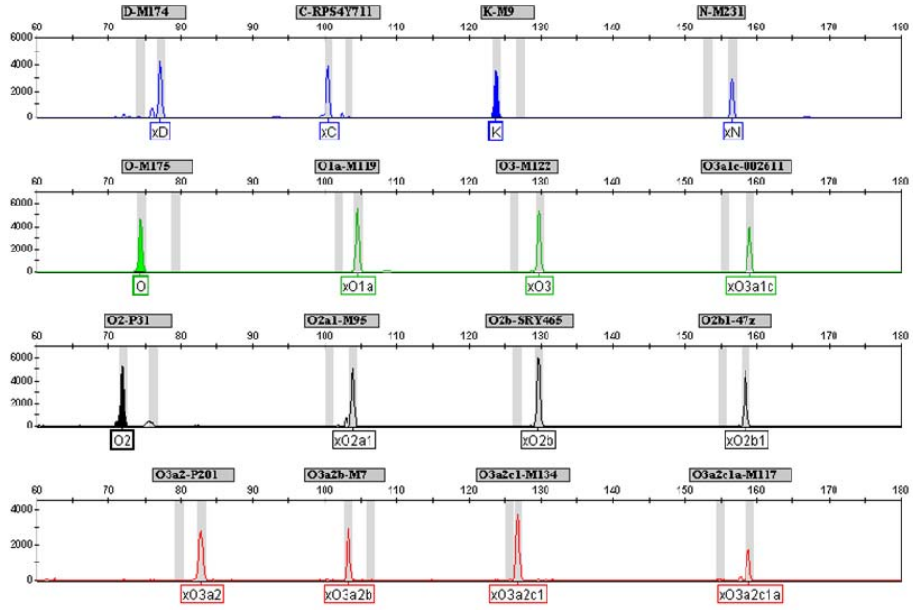
Haplogroup O3

g



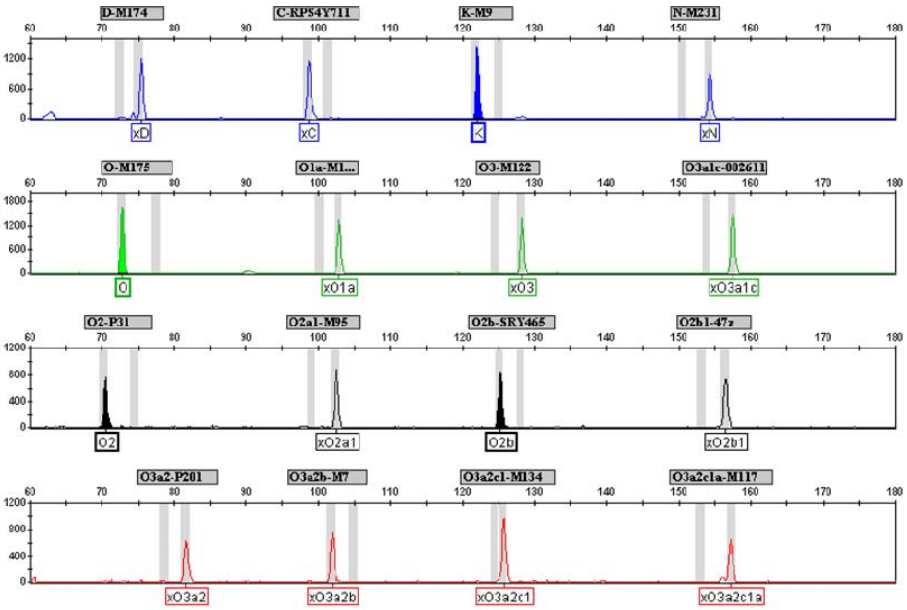
Haplogroup O3a1

h



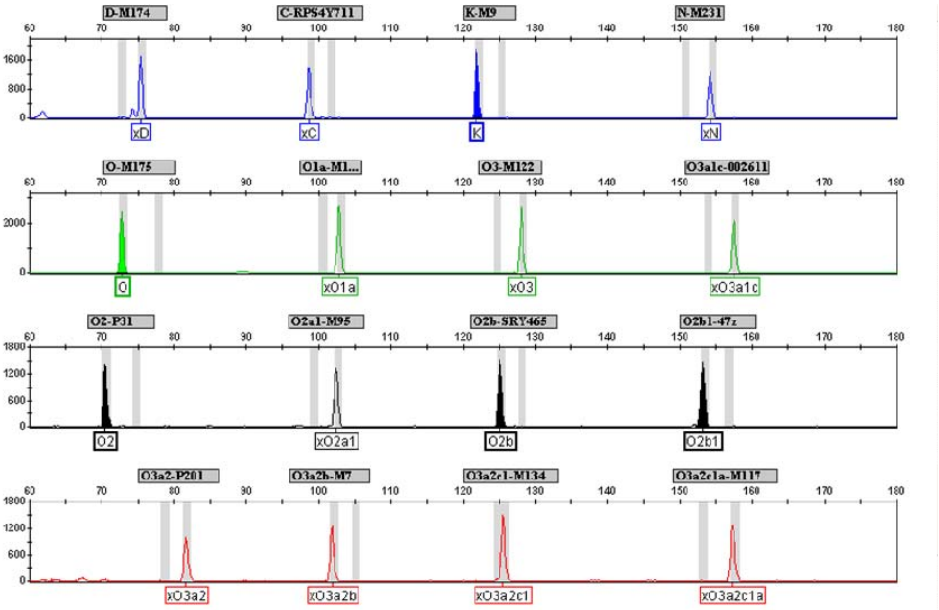
Haplogroup O2

i



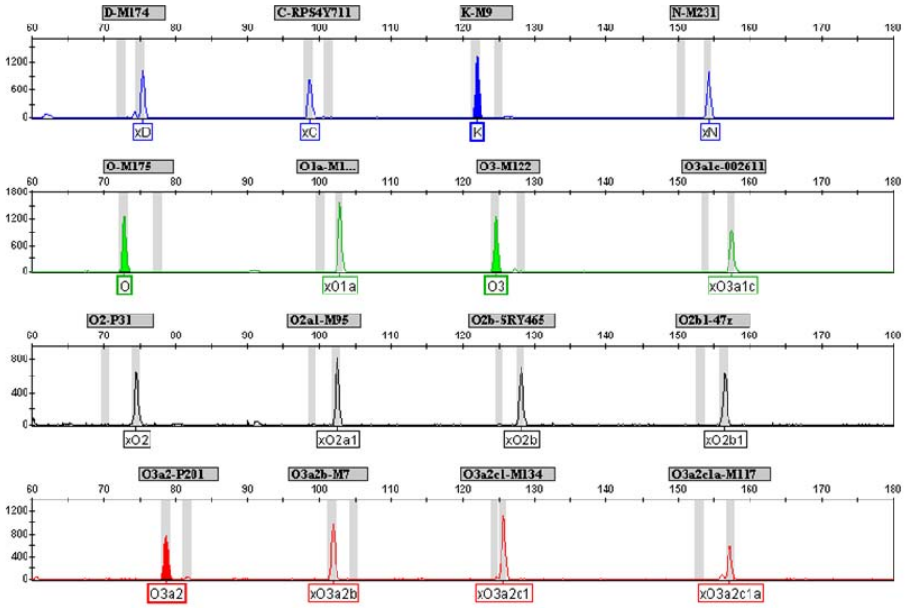
Haplogroup O2b

j



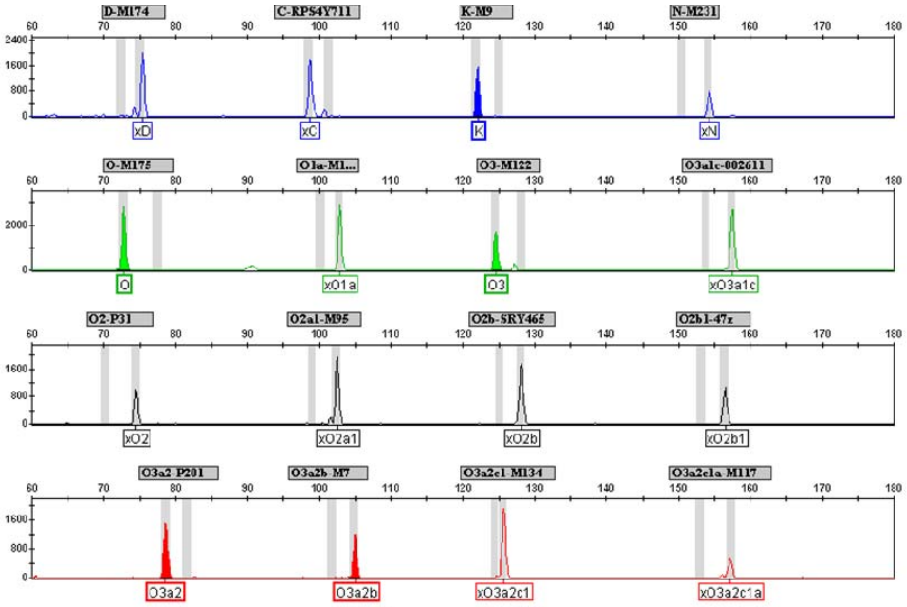
Haplogroup O2b1

k



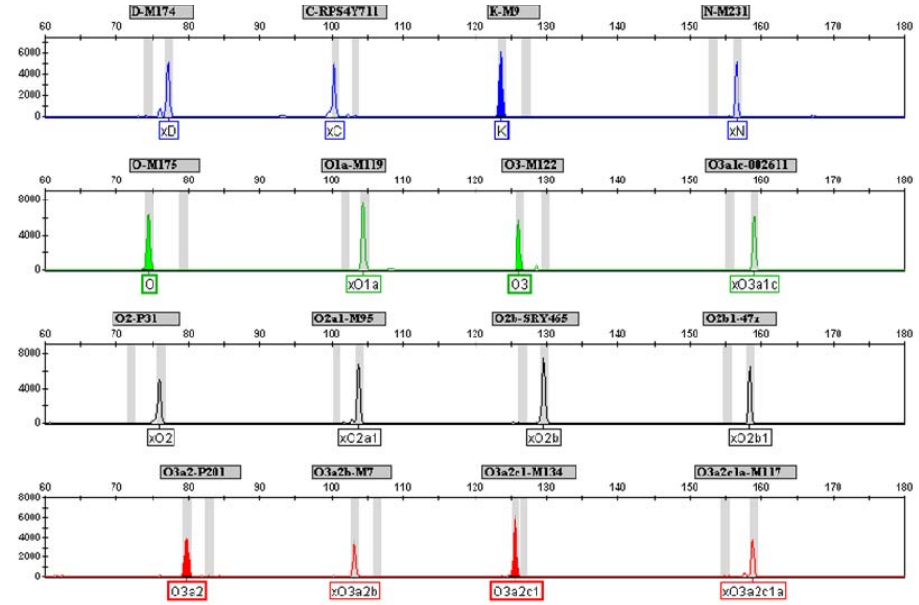
Haplogroup O3a2

l



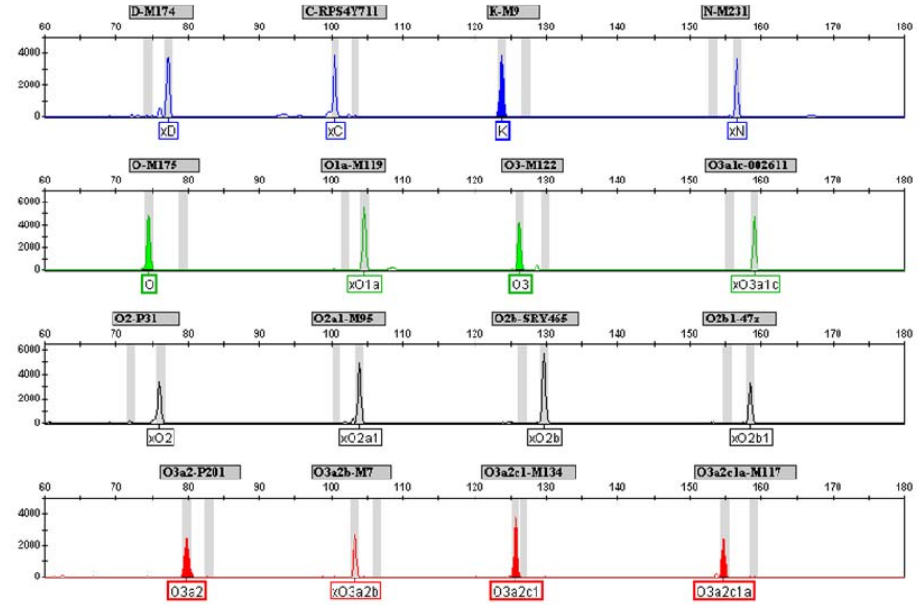
Haplogroup O3a2b

m



Haplogroup O3a2c1

n



Haplogroup O3a2c1a

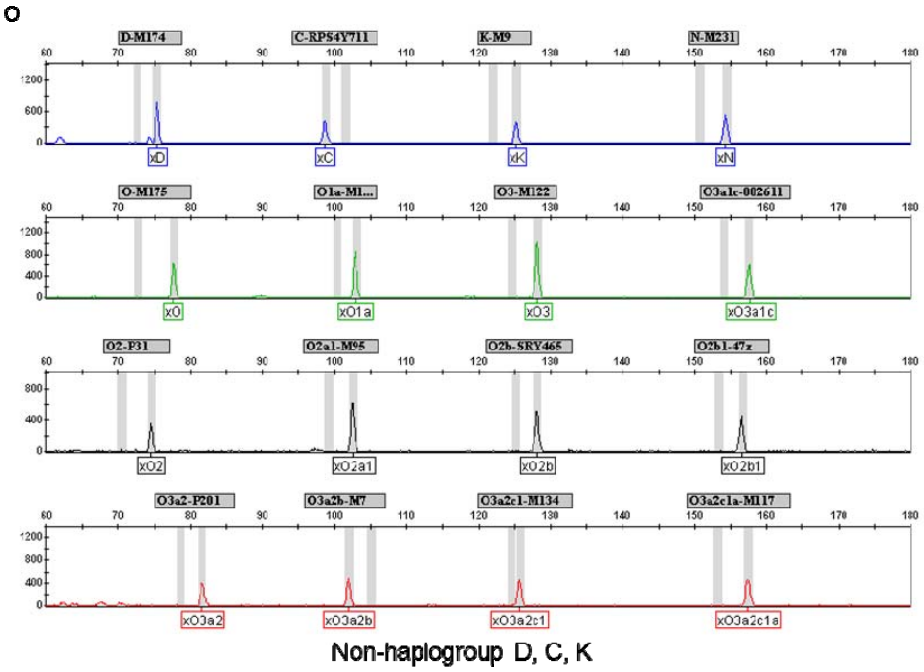


Figure 7. Representative electropherograms for multiplex AS-PCR assay in Korean population. Haplogroup D-M174 (a) Haplogroup C-RPS4Y$_{711}$ (b) Haplogroup K-M9 (c) Haplogroup N-M231 (d) Haplogroup O1a-M119 (e) Haplogroup O3-M122 (f) Haplogroup O3a1c-JST002611 (g) Haplogroup O2-P31(h) Haplogroup O2b-SRY$_{465}$ (i) Haplogroup O2b1-47z (j) Haplogroup O3a2-P201 (k) Haplogroup O3a2b-M7 (l) Haplogroup O3a2c1-M134 (m) Haplogroup O3a2c1a-M117 and (n) Non-haplogroup D-M714, C-RPS4Y$_{711}$ and K-M9 (o).

C. Concordance between the two developed multiplex PCR systems

The performance of the two developed multiplex PCR systems was assessed and compared by analyzing 300 DNA samples. Full concordance for determination of haplogroups between the multiplex PCR systems was observed except for one sample with M117 but without M133 mutation which were both reported to designate haplogroup O3a2c1a (Figure 8). Sequence structures for

the two mutations were confirmed by direct sequencing analysis, which showed the possibility for non-equivalence between M133 and M117 (Figure 9).

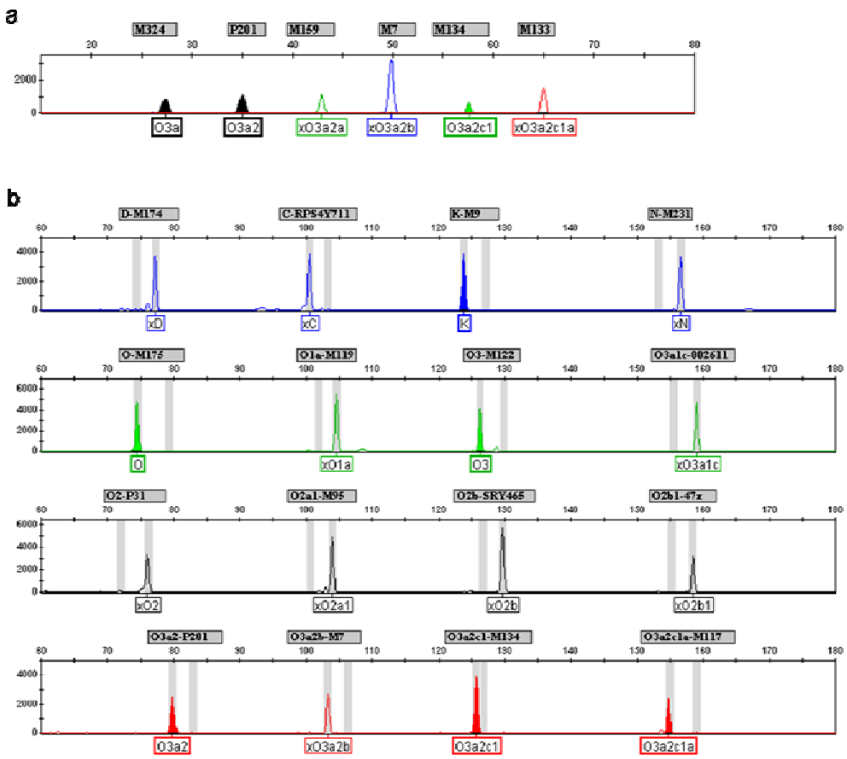


Figure 8. Electropherograms of a sample which seems belonging to haplogroup O3a2c1 using multiplex SBE-III but seems belonging to haplogroup O3a2c1a using multiplex AS-I.

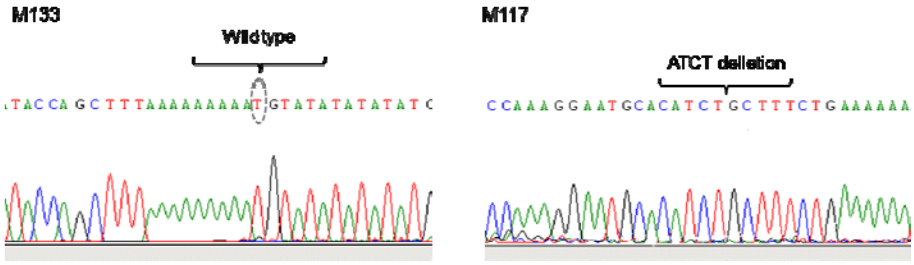


Figure 9. Sequence analysis of a sample which shows discordance between Y-SNPs M133 and M117 designating haplogroup O3a2c1a.

D. Sensitivity of the developed multiplex PCR systems

To evaluate the newly developed multiplexes, their sensitivities were tested with five replicates of DNAs at various concentrations. For the multiplex PCRs followed by SBE reactions, average peak height from five replicates showed no linear correlation with DNA concentrations under the condition of 33 PCR cycles (Figure 10a). This was explained by the fact that the amount of amplified products in plateau phase after approximately 30 cycles was not proportional to the amount of input DNA.[47] When the increased PCR cycles were used for the diluted DNA samples to improve the amplification yield (Figure 11a), the peak intensities were increased and allele losses showed a tendency to decrease (Figure 10b vs. 11b). However, the less input DNA used, the more standard deviation increased, which was caused by the stochastic effect.[48] As a result, all 27 SNPs were successfully typed with DNA amounts as low as 62 pg and 35 PCR cycles, at which only P31 marker failed to generate signal once (Figure 11b). For 31 pg of template DNA with 37 PCR cycles, allele losses were observed at $RPS4Y_{711}$ and M119 markers in one of five (20%) and at M214, M175, M7, M159, and M217 markers in two of five (40%). For 15 pg of template DNA, most markers showed allele losses except for the $RPS4Y_{711}$ and M407 markers with 37 PCR cycles. The alleles of M89, M175, M7 and JST002611 markers were most frequently lost (80%). Based on these results, these multiplex SBE reactions were proven to be able to successfully type 1 ng

to 62 pg of DNA, and multiplex SBE-VI was the most sensitive SBE reaction (average loss; 4.3%), followed by multiplex SBE-V (5.7%), II (6.2%), III (8.1%), IV (9.7%) and I (10.0%).

On the other hand, the reliable results for 19 Y-SNPs were showed in DNA amounts as low as 250 pg without allelic drop in or drop-out using the multiplex AS-PCR assays (Figure 12). To enhance the sensitivity for the Multiplexes, increased PCR cycles up to 35 were used. However, pseudo-positive peaks were appeared, which might be caused by decreasing specificity of the allele-specific primers (Figure 13). Thereby, the 30 PCR cycles were proper to amplify the multiplex AS-PCR assays. For 125 pg of template DNA, allele losses were appeared at M7, M117 and SRY_{465} in one of five (20%), at M134 in two of five (40%) and at P31 in four of five (80%), which markers are all included in multiplex AS-I. For 62 pg of template DNA, most markers consisted of multiplex AS-I showed allele losses except M174, M119 and M122 whereas allele losses were observed only at occasion at P164 and M134 in multiplex AS-II. As a result, multiplex AS-II showed more sensitive result than multiplex AS-I. Based on these results, the multiplex SBE reactions are more suitable to analyze low concentration of DNA than the multiplex AS-PCR assays.

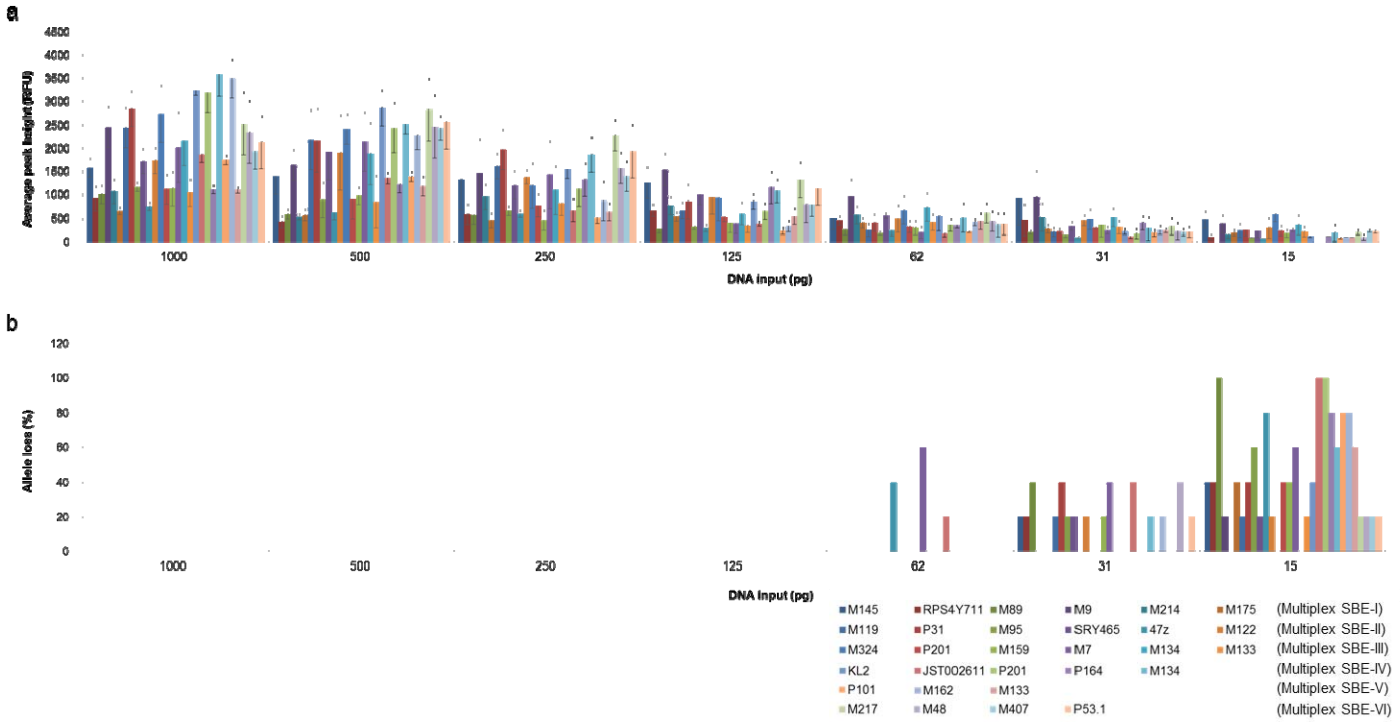


Figure 10. Sensitivity tests for multiplex SBE-I, II, III, IV, V and VI with 9948 male or 2800M control DNA at various concentrations under the same PCR conditions. Average peak heights were calculated from five replicates and error bar indicates the standard deviation. (a) Percentages of observed allele loss are presented at each DNA concentration. (b)

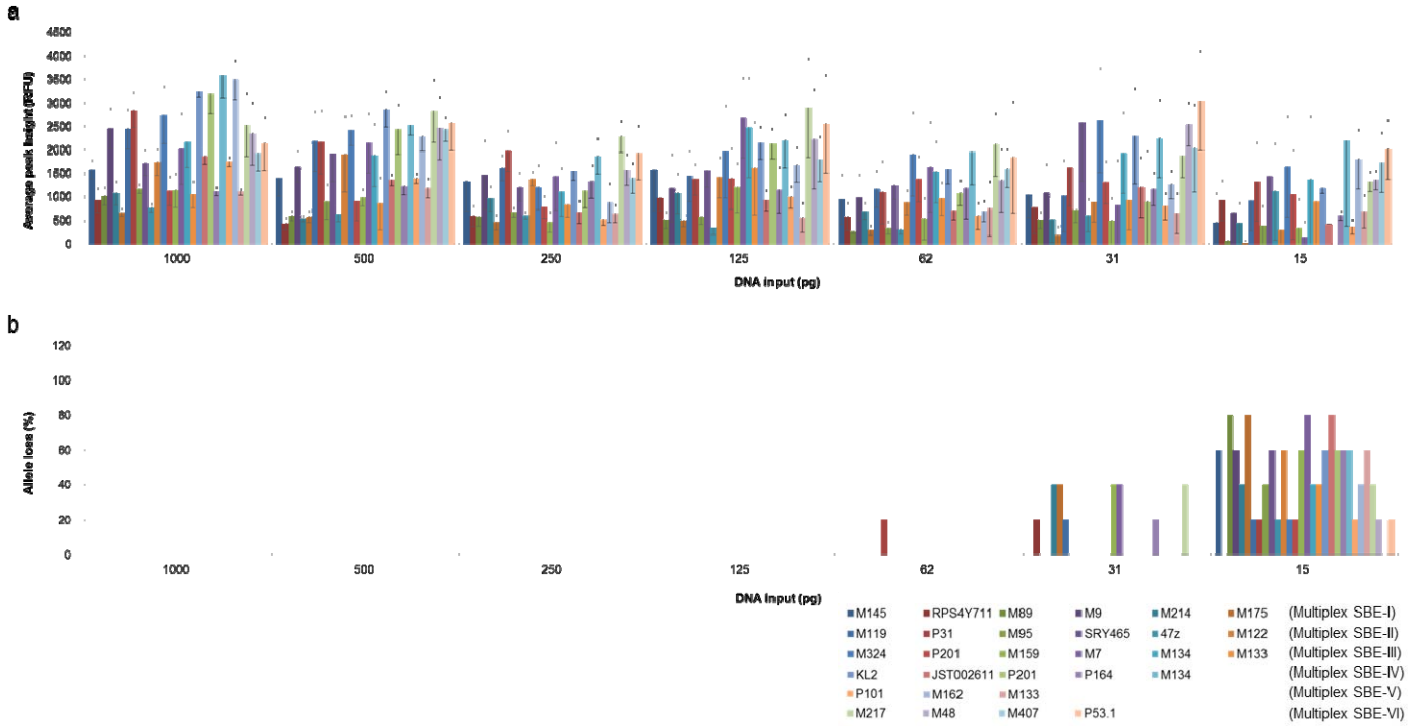


Figure 11. Sensitivity tests for multiplex SBE-I, II, III, IV, V and VI with 9948 male or 2800M control DNA using different PCR cycles depending on DNA concentraction: 33 cycles for 1000, 500 and 250 pg of template DNA; 35 cycles for 125 and 62 pg; and 37 cycles for 31 and 15 pg. Average peak heights were calculated from five replicates and error bar indicates the standard deviation. (a) Percentages of observed allele loss are presented at each DNA concentration. (b)

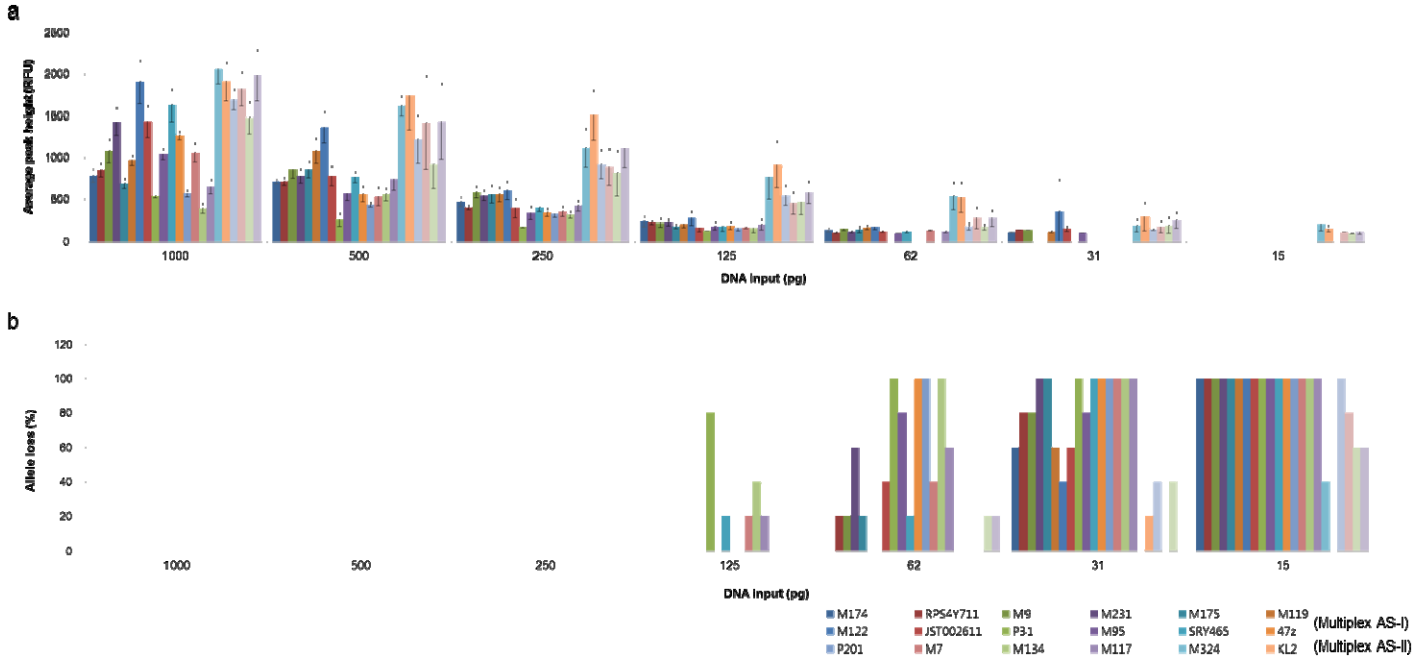


Figure 12. Sensitivity tests for multiplex AS-I and II with 2800M control DNA sample at various concentrations under the same PCR conditions. Average peak heights were calculated from five replicates and error bar indicates the standard deviation. (a) Percentages of observed allele loss are presented according to the DNA concentrations. (b)

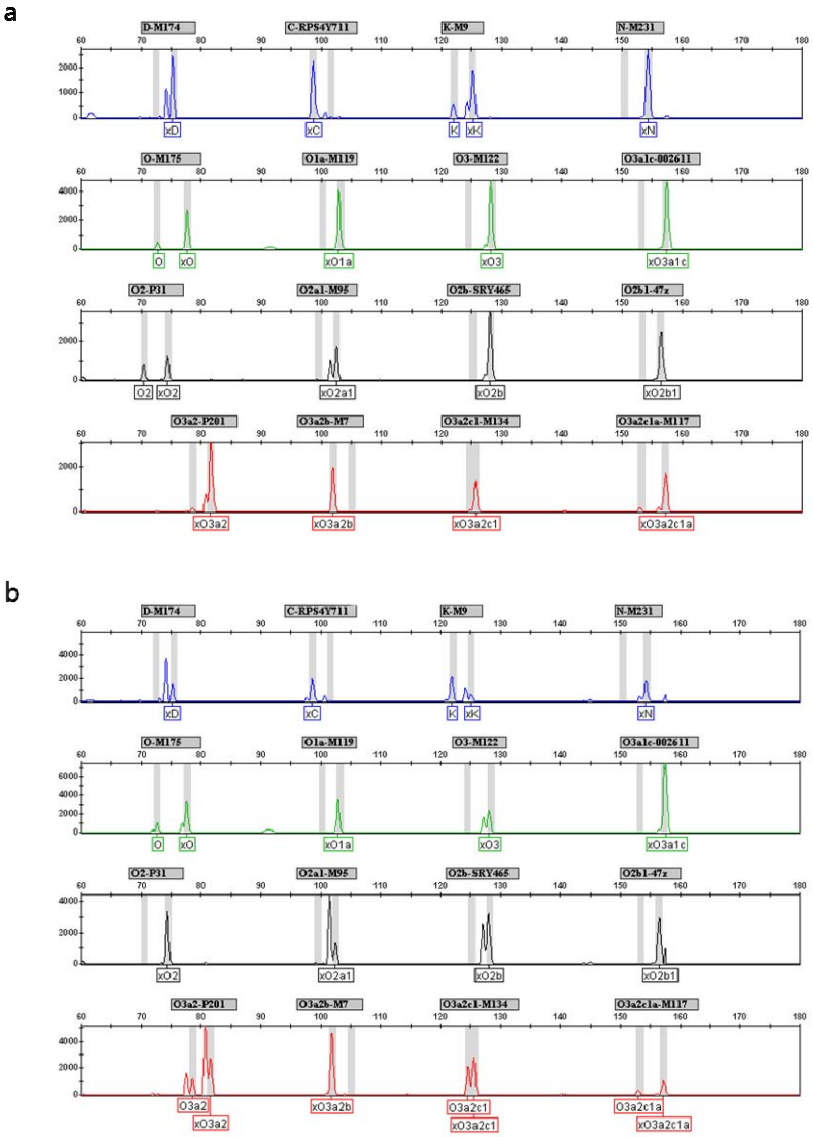


Figure 13. Representative electropherograms of 9948 male DNA at 125 pg using 33 cycles of multiplex AS-I (a) and at 31 pg with 35 cycles (b), respectively.

E. Efficiency of the developed multiplex PCR systems

To test the efficiency of the multiplex systems, the Y-SNPs were also typed in one artificially degraded DNA sample. For the multiplex PCRs

followed by SBE reaction, the results from the replicate analyses of the artificially degraded DNA sample and its intact DNA sample showed full allele concordance although the signal intensities for all SNP alleles were decreased by a percentage ranging from 58.9% to 93.6%, (Figs. 14a and 15a). For the artificially degraded DNA sample, peak intensities among the loci lost the balance due to poor amplifications of some markers such as $RPS4Y_{711}$, M214, SRY_{465}, and M48. The correct allelic determinations with no dropout or drop-in and small values of standard deviation from average peak height between independent replicates demonstrate that the multiplex SBE reactions are reliable and reproducible. Whereas, for the multiplex AS-PCR assays, some allelic genotypes could not be typed in the artificially degraded DNA because of no amplification at M231, M119, P31, SRY_{465}, 47z, M117 and JST002611 and failure of allele-specific primers to discriminate between the allele state (Figs. 14b, 15b and 15c). This result showed the multiplex SBE reactions are well-suited for the analysis of degraded DNA samples.

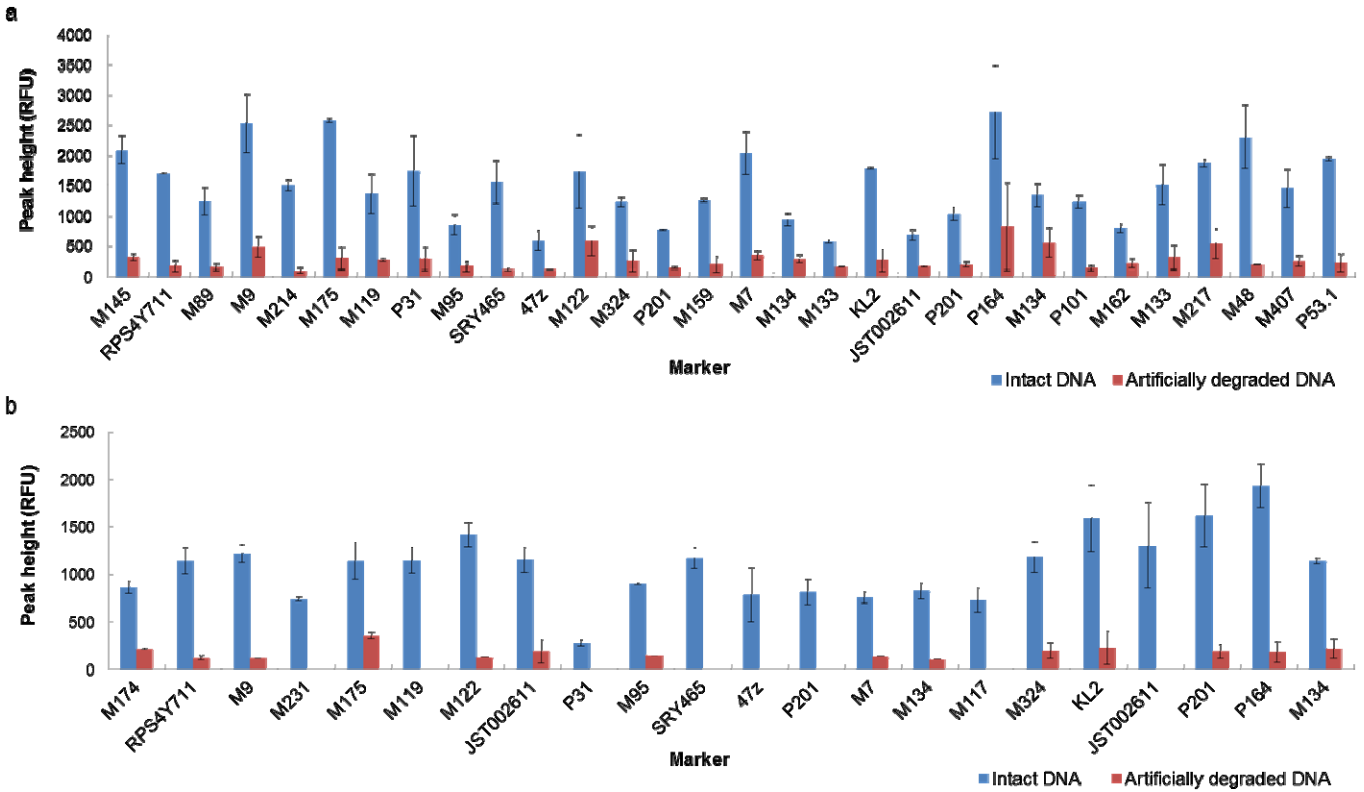


Figure 14. Peak height comparison between intact and artificially degraded DNA at each Y-SNP locus using multiplex SBE reactions (a) and using multiplex AS-PCR assays (b), respectively. Intact and artificially degraded DNA was amplified in replicates and values are expressed at the average peak height at each locus. The error bar indicates the standard deviation from the average.

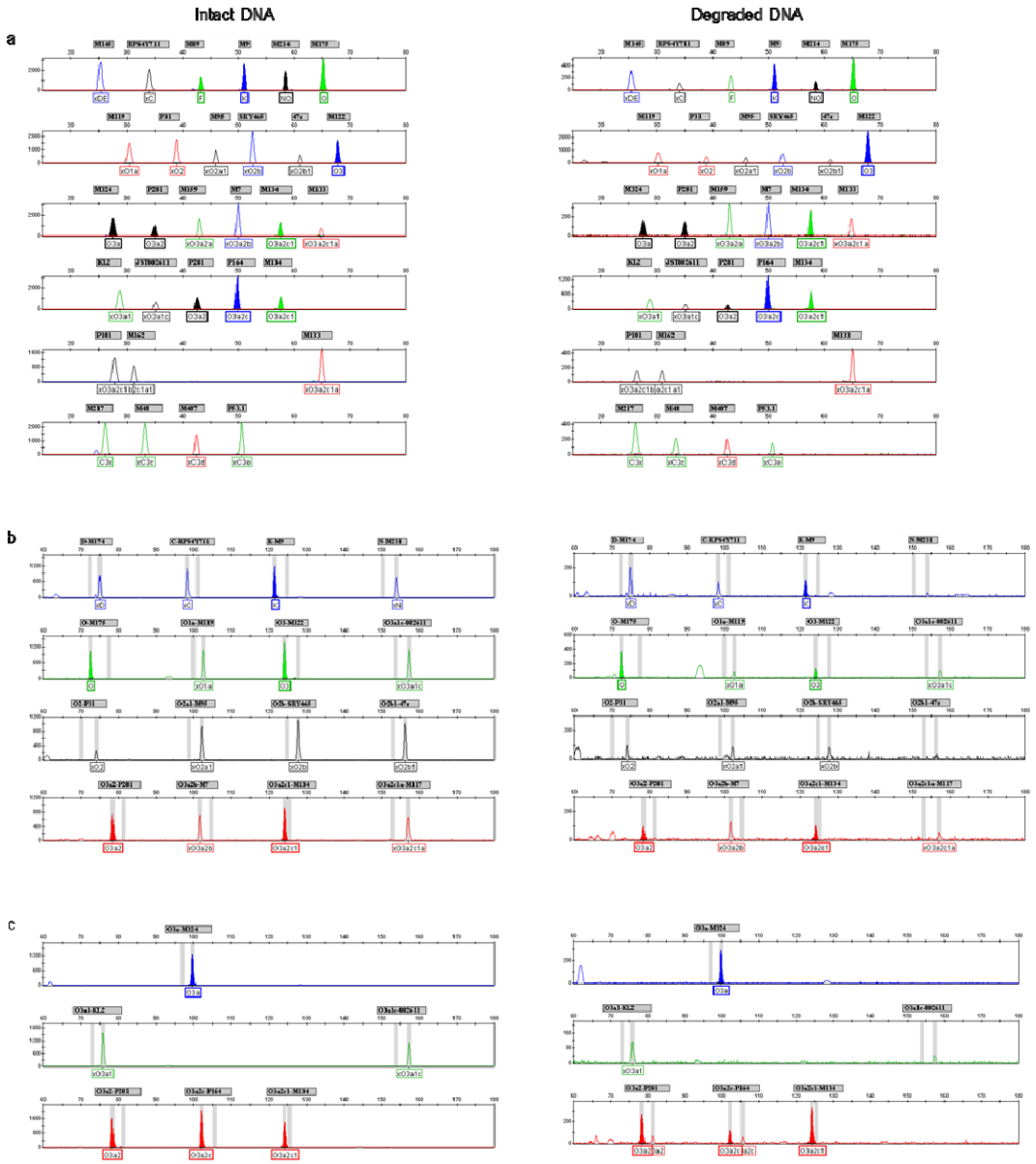


Figure 15. Representative electropherograms of intact and artificially degraded DNA using multiplex SBE reactions I-VI from top to bottom (a), multiplex AS-I (b) and multiplex AS-II (c), which belong to haplogroup O3a2c1.

Thereby, the multiplex SBE reactions were further tested with the 10 DNA samples extracted from 55-year-old skeletal remains and could assign the samples into the relevant haplogroups (Figure 16 and Table 7). For amplification of the DNA samples from the skeletal remains, a higher PCR cycle number was used to enhance amplification efficiency. No allele dropout was observed in the 10 DNA samples from the skeletal remains using 35 PCR cycles and as little as 62 pg of template DNA. Allele drop-in was observed in some markers such as $RPS4Y_{711}$, M214, P31, and SRY_{465}, but allele drop-in interfering with determination of relevant haplogroups was not observed in the consensus profiles obtained by duplicate amplification. Although a single incident of allele drop-in was detected in the negative controls and reagent blanks, they were not reproduced upon re-amplification. In addition, when comparing the genotyping success of Y-SNP markers with those of STR markers in the same samples from a previous study[35], the Y-SNP markers provided higher success than STR markers, especially in lower quality DNA samples (Table 7). This result suggests that the multiplex SBE reactions may be very useful for analyzing degraded forensic and ancient samples if the consensus profiles obtained from replication of analysis were used with negative control and reagent blank to illustrate that there has been a source of contamination in the analytical method.[39, 40]

Table 7. Y haplogroup affiliations, success rates of STR typing and DNA concentrations of 10 old skeletal remains

Sample	Y haplogroup	Success of STR typing[a] (%)	Concentration[a] (pg/μl)
1	NO	93.3	114.8 ± 10.20
2	O2*	100	205.7 ± 2.75
3	O2b*	66.7	55.9 ± 5.94
4	O2b*	100	106.5 ± 3.56
5	O3a2*	100	766.1 ± 39.03
6	O3a2c1	80.0	149.8 ± 11.30
7	O1a	33.3	27.8 ± 0.13
8	O2b*	100	169.9 ± 10.96
9	O3a2*	100	84.9 ± 14.71
10	O2b*	100	275.2 ± 57.04

[a] Success of STR typing using the AmpFlSTR® Identifiler® PCR Amplification kit and DNA concentrations reported in a previous study[35].

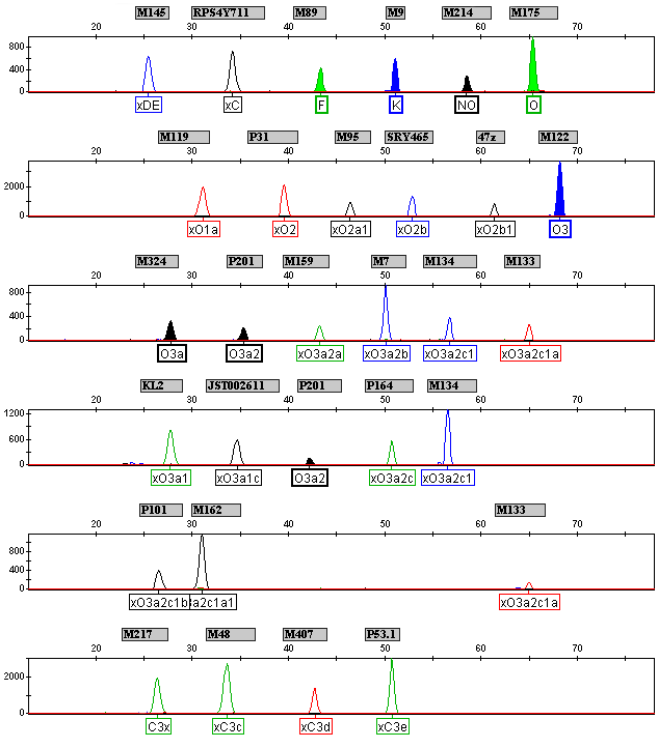


Figure 16. Representative electropherograms of multiplex SBE reactions I-VI from top to bottom with a DNA sample from 55-year-old skeletal remains which belong to haplogroup O3a2*.

2. Y chromosome haplogrouping of Koreans

DNA samples from 1006 Korean males were analyzed using the multiplex allele-specific PCR assays. A set of 16 Y-SNPs (M7, M9, M95, M117, M119, M134, M174, M175, M122, M231, P31, P201, JST002611, RPS4Y$_{711}$, SRY$_{465}$ and 47z) was initially analyzed in all samples using multiplex AS-I to determine the haplogroups frequent in East Asians. According to the results obtained from multiplex AS-I, the samples belonging to haplogroup O3-M122 were then typed using multiplex AS-II to further divide the subhaplogroups O3 according to the revised tree of haplogroup O[30]. In addition, to detect rare haplogroups and to confirm the position of the newly updated markers (JST002611, KL2, P164 and PK4), subsequent typing of the remaining samples was performed hierarchically using a monoplex SBE reaction depending on the haplogroup designation results from multiplex AS-II. Representative electropherograms of the monoplex SBE reactions are shown in Figure 17. A total of 21 different haplogroups were identified (when M117 was used alone to designating haplogroup O3a2c1a); haplogroup O2b*-SRY$_{465}$ was most frequently observed (24.2%), followed by haplogroups C3-M217 (xC3c, C3d, C3e) (13.4%) and O3a2c1a-M117 (12.3%) (Figure 18). The haplogroup diversity was 0.8830, and the discriminatory capacity was 2.1%.

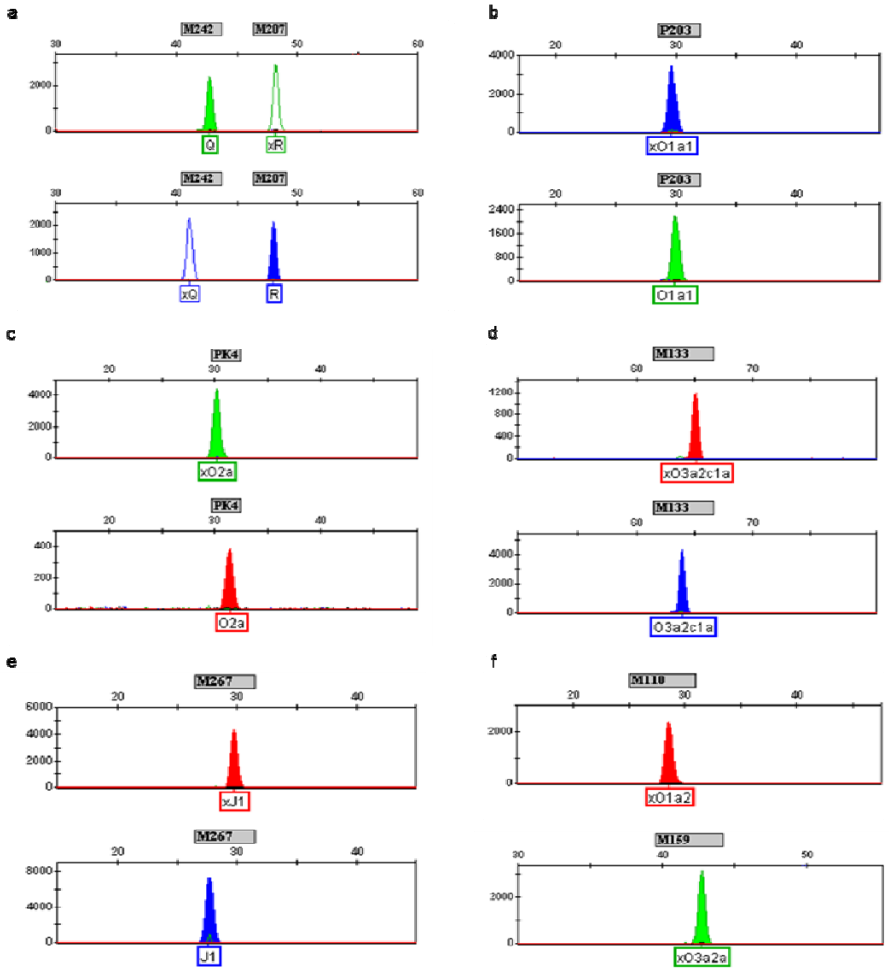


Figure 17. Representative electropherograms of monoplex SBE reactions. Haplogroup Q-M242 (upper) and haplogroup R-M207 (lower), respectively (a) Ancestral (upper) and derived (lower) allelic state at P203, which defines haplogroup O1a1(b) Ancestral (upper) and derived (lower) allelic state at PK4, which defines haplogroup O2a (c) Ancestral (upper) and derived (lower) allelic state at M133, which defines haplogroup O3a2c1a (d) Ancestral (upper) and derived (lower) allelic state at M267, which defines haplogroup J1 (e) Ancestral allelic state at M110 (upper) and M159 (lower), which define haplogroups O1a2 and O3a2a, respectively (f).

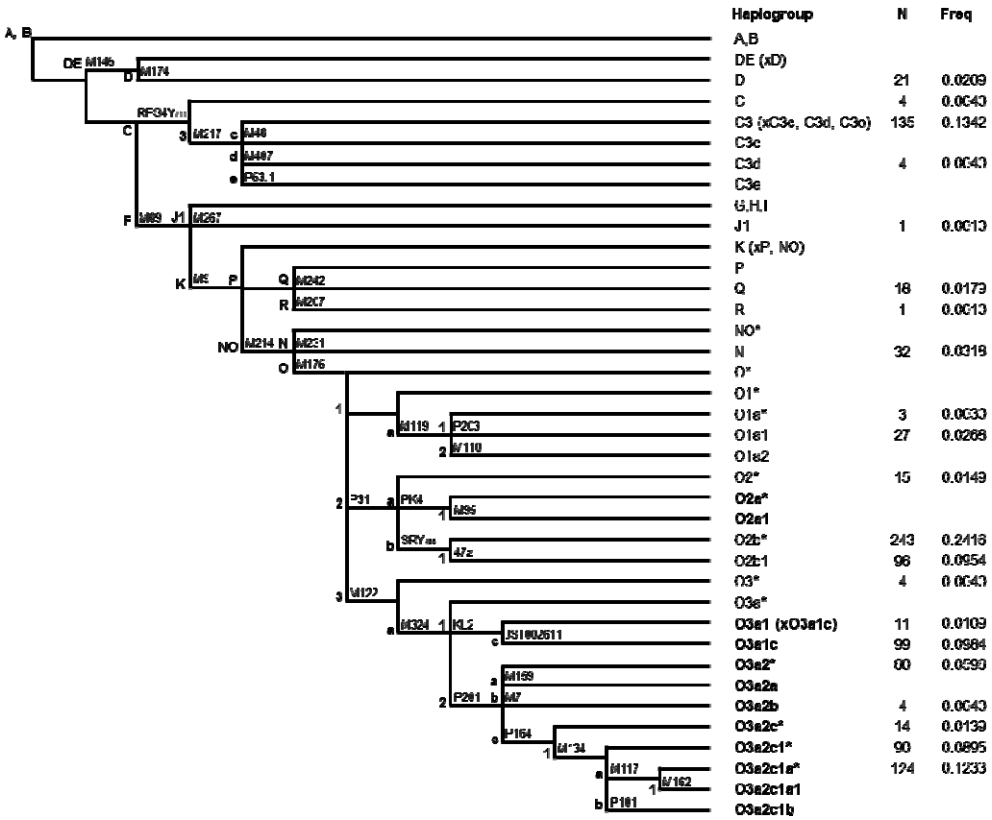


Figure 18. Phylogenetic tree of the 35 Y-chromosomal binary polymorphisms analyzed in 1006 Korean males. The analyzed Y-SNPs are shown in each branch, and the corresponding haplogroups and observed frequencies are shown at the end of each branch according to Karafet et al.[29] and Yan et al.[30] The lineages renamed by Yan et al.[30] are indicated in *bold*.

By typing of the newly defined or relocated SNPs KL2, JST002611, P164 and PK4 in Korean haplogroup O-M175 samples, their phylogenetic positions were confirmed. The KL2 mutation was found in all samples belonging to O3a-M324 (xP201). Moreover, the chromosomes with the KL2 mutation were divided into two groups based on the presence or absence of the JST002611 mutation. On the other hand, none of the samples shared mutations KL2 and

P201, so their relative position could be determined. In addition, the mutation P164 was observed in all samples with the M134 mutation but was not found in the samples with the M7 mutation. I also found that some samples belonging to O3a2-P201 (xM159, M7, M134) had the P164 mutation, thereby corresponding to paragroup O3a2c*-P164. Therefore, the positions of SNPs KL2, JST002611 and P164 could be confirmed in these surveyed samples. However, the relative position of the PK4 marker could not be determined because of a lack of positive samples.

The addition of KL2 and P164 markers to the phylogeny of haplogroup O improved the resolution of the Korean haplogroup O3a-M324, which had the same result as the Han Chinese population.[30] However, the distribution of the subhaplogroups differed significantly from those in the East, North and South Han Chinese populations[30] (p<0.0001). Specially, there were significant differences in the distributions of sublineages inside haplogroups O1 and O2. While haplogroup O1a1-P203 was present in only 2.7 % of Koreans, it was found in 13.0% of the Han Chinese population.[30] Haplogroup O2-P31, another derived sublineage of O-M175, was divided into subgroup with PK4 mutation and subgroup with SRY_{465} mutation. Haplogroup O2a defined by revised phylogenetic position of PK4 in the updated tree of haplogroup O was absent in the Korean population, but is abundant in the South Han Chinese population. Meanwhile, O2b-SRY_{465} and its derived sublineage O2b1-47z were found frequent (37.7%) in the Korean population, but they were nearly absent in the

Han Chinese population. The subhaplogroups were concentrated in Korean and Japanese populations[32] and the results of this current study are consistent with those findings.

3. Relevance of combined haplogroup and haplotype analyses

A. Non-equivalence between M117 and M133 markers

Combined haplogroup and haplotype analyses were performed for 706 DNA samples with genotyped 22 Y-STR loci. Unexpectedly, I found that the M117 marker was not always observed phylogenetically equivalent to the M133 when the multiplex allele-specific PCR results were compared to those of SBE reaction (Figure 8). All but eight of the analyzed samples with the M117 mutation had the M133 mutation (M117+, M133+), whereas the eight samples were detected without the M133 mutation (M117+, M133-). Sequence structures of the eight samples were also confirmed by direct sequencing analysis.

In order to evaluate the relationships of these samples (M117+, M133-) with those of the other samples (M117+/M133+ and M117-/M133-) within haplogroup O3a2c1-M134, a network was constructed using Y-STR haplotypes (Figure 19). Haplotypes produced two major clusters, clearly separated by the presence or absence of both

M117 and M133 mutations. A cluster formed by the haplotypes with (M117+, M133-) mutation was distinct from the major cluster formed by the haplotypes with (M117+, M133+) mutation, which indicated non-equivalence of M133 to M117.

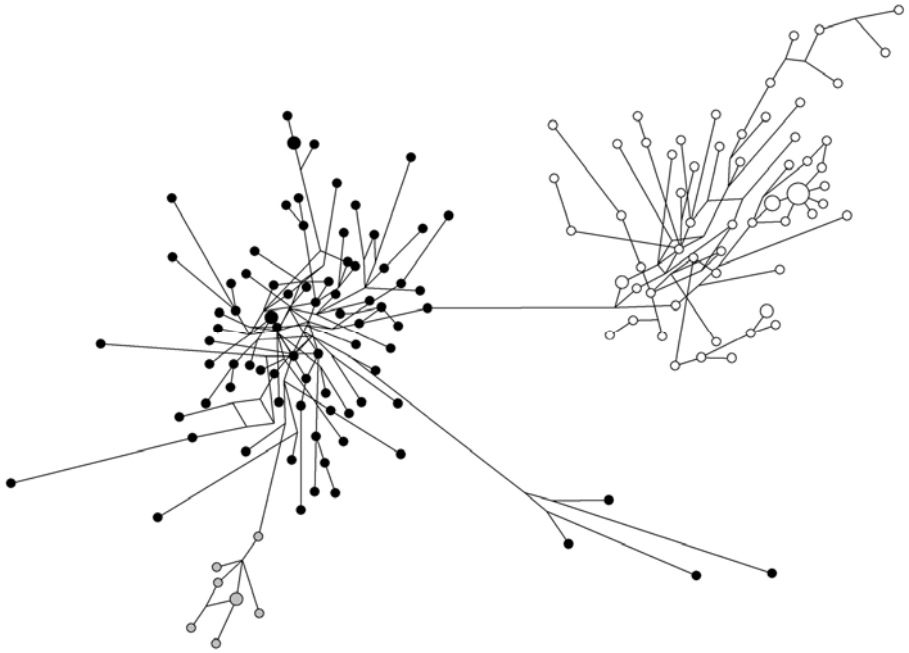


Figure 19. Relationships between Y haplotypes within haplogroup O3a2c1-M134. The median-joining network was based on information from 17 Y-STR loci (DYS19, DYS388, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS446, DYS447, DYS448, DYS449, DYS456, DYS458, DYS635 and GATA H4.1). The haplotypes carrying both M117 and M133 mutations are indicated by filled circles. The haplotype carrying neither M117 nor M133 mutations is indicated by an open circle. The haplotypes carrying the M117 mutation, but not M133 mutation, are indicated by gray circles.

B. Y-STR markers and their relationships with haplogroups

(A) Variability of Y-STR markers and its relationship with haplogroup

Allele frequencies and gene diversities at each single copy Y-STR locus are shown in Table 8. Haplotype frequencies and diversity values at each multi-copy locus are shown in Table 9. In 706 Korean population samples, almost all Y-STRs reached the gene diversity value of >0.5 except for DYS388 (0.4850), DYS391 (0.2667) and DYS437 (0.4320). In particular, the diversities were high in order of DYS464, DYS385, DYS449, and DYS446 in a Korean population.

To assess how much of Y-STR variability is partitioned by difference among haplogroup background and which Y-STR contributes more to the difference among haplogroups, a separate AMOVA analysis was performed based on Y-STR information (Table 10). DYS392, DYS393, DYS437, DYS448 and DYS388 markers showed more than 60% of their total genetic variation among haplogroups, and these markers have relatively slow mutation rates than others.

Table 8. Allele frequencies and gene diversity of 19 Y-STR loci in a Korean population

Allele	DYS 19	DYS 388	DYS 389I	DYS 389II	DYS 390	DYS 391	DYS 392	DYS 393	DYS 437	DYS 438	DYS 439	DYS 446	DYS 447	DYS 448	DYS 449	DYS 456	DYS 458	DYS 635	GATA H4.1
7						0.0014													
8																			
9						0.0269				0.0085	0.0028								
10		0.1275				0.8470	0.0028	0.0014		0.5043	0.0439								
11		0.0028	0.0113			0.1246	0.1544			0.1530	0.2861	0.0751							
12		0.6856	0.3725				0.1020	0.3810		0.0269	0.5156	0.1346				0.0170			
13	0.0241	0.1714	0.2337				0.4901	0.4632	0.0014	0.2989	0.1346	0.3584				0.0142	0.0085		
14	0.1544	0.0085	0.3739				0.2167	0.1062	0.6941	0.0085	0.0170	0.2181				0.1317	0.0071		
15	0.3810	0.0043	0.0071				0.0241	0.0425	0.2946			0.0779				0.6756	0.0722		
16	0.3286		0.0014				0.0099	0.0057	0.0099			0.0680				0.1204	0.1516		
17	0.1091											0.0411		0.0340		0.0312	0.3102		
18	0.0014											0.0170	0.0028	0.3569		0.0099	0.2833		
19												0.0099	0.0057	0.2380			0.1331	0.0524	0.0425
20					0.0014									0.2337			0.0255	0.2734	0.4278
21					0.0057								0.0014	0.0963			0.0057	0.4660	0.4448
22					0.1232								0.0071	0.0298				0.1331	0.0822
23					0.4873								0.1459	0.0028				0.0467	0.0028
24					0.2436								0.2550					0.0269	
25[a]					0.1246								0.3810		0.0028			0.0014	
26				0.0014	0.0113								0.1190		0.0043				
27				0.0737	0.0014								0.0538		0.0326				
28				0.2408									0.0213		0.0524				
29				0.3980									0.0057		0.1147				
30				0.2280											0.2196				
31				0.0524											0.2096				
32				0.0057											0.1601				
33															0.0963				
34															0.0666				
35															0.0156				
36															0.0028				
42[*]															0.0014				

Table 8. Cont.

Allele	DYS 19	DYS 388	DYS 389I	DYS 389II	DYS 390	DYS 391	DYS 392	DYS 393	DYS 437	DYS 438	DYS 439	DYS 446	DYS 447	DYS 448	DYS 449	DYS 456	DYS 458	DYS 635	GATA H4.1
Microvariants, duplicated and null alleles																			
14.1																	0.0014		
17.2																	0.0014		
27.2															0.0014				
28.2															0.0014				
29.2															0.0014				
30.1[*]															0.0014				
30.2															0.0170				
16-17	0.0014																		
22-23					0.0014														
23-24													0.0014						
Null														0.0085					
h	0.7115	0.4850	0.6675	0.7244	0.6733	0.2667	0.6789	0.6281	0.4320	0.6330	0.6329	0.7886	0.7520	0.7510	0.8522	0.5108	0.7779	0.6857	0.6115

[a] At DYS449, these alleles determined by amplicon size (i.e. ostensible allele) and only one of allele 25 was ostensible allele.

Table 9. Haplotype distribution and diversity values in multi-copy Y-STRs

DYS385				DYS464			
Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency	Haplotype	Frequency
8-20 [a]	0.0014	12-21	0.0014	12	0.0014	13-15-18	0.0128
9-16 [a]	0.0014	13-13	0.0227	13	0.0128	13-15-19	0.0014
9-18 [a]	0.0028	13-14	0.0071	14	0.0028	13-16-17	0.0283
9-19	0.0156	13-15	0.0028	15	0.0028	13-16-18	0.0113
9-20	0.0028	13-16	0.0170	16	0.0085	13-17-18	0.0028
10-10	0.0014	13-17 [a]	0.0213	17	0.0128	14-15-16	0.0043
10-16	0.0043	13-18 [a]	0.0538	11-17[b]	0.0014	14-15-17	0.0028
10-17	0.0482	13-19	0.0382	12-13	0.0043	14-15-18	0.0014
10-18	0.1048	13-20	0.0255	12-14	0.0014	14-16-17	0.0071
10-19 [a]	0.0977	13-21	0.0028	12-16	0.0057	14-16-18	0.0028
10-20	0.0340	14-14	0.0028	13-14	0.0467	15-16-17	0.0014
10-21	0.0113	14-16	0.0028	13-15	0.0453	11-12-13-15	0.0014
11-11	0.0113	14-17	0.0128	13-16	0.0128	11-12-13-16	0.0014
11-12	0.0184	14-18	0.0213	13-17	0.0057	11-13-14-16	0.0014
11-13	0.0142	14-19	0.0113	13-18	0.0014	11-13-16-18	0.0014
11-14	0.0071	14-20	0.0156	14-15	0.0057	11-15-16-17	0.0014
11-15	0.0014	14-21	0.0085	14-16[b]	0.0142	12-13-14-15	0.0354
11-16	0.0128	14-22	0.0043	14-17	0.0014	12-13-14-16	0.0312
11-17	0.0283	15-15	0.0028	15-16	0.0198	12-13-14-17	0.0113
11-18	0.0099	15-17	0.0014	15-17	0.0071	12-13-15-16	0.0113
11-19	0.0354	15-18	0.0014	16-17	0.0113	12-13-15-17	0.0057
11-20	0.0156	15-19	0.0071	16-18	0.0028	12-13-15-18	0.0014
11-21	0.0028	15-20	0.0184	17-18	0.0028	12-13-16-17	0.0028
12-12	0.0028	15-21	0.0128	14-14.3	0.0071	12-13-16-18	0.0043
12-13	0.0043	15-22	0.0085	10-13-15	0.0014	12-13-16-19	0.0014
12-14	0.0113	15-23	0.0014	10-14-16	0.0014	12-14-15-16	0.0298
12-15	0.0043	16-17	0.0014	11-12-17	0.0043	12-14-15-17	0.0057
12-16[a]	0.0241	16-20	0.0028	11-13-14	0.0014	12-14-15-18	0.0014
12-17	0.0382	16-21	0.0028	11-13-15	0.0014	12-14-16-17	0.0028
12-18	0.0453	17-20	0.0014	11-14-16	0.0014	12-14-16-18	0.0014
12-19	0.0411	19-19	0.0014	12-13-14	0.0255	12-15-16-17	0.0071
12-20	0.0085	19-20	0.0014	12-13-15	0.0467	12-15-16-18	0.0014
				12-13-16	0.0043	12-15-17-18	0.0014
				12-13-17	0.0014	12.3-14-15-17	0.0014
				12-14-15	0.0255	13-14-15-16	0.0128
				12-14-16	0.0340	13-14-15-17	0.0099
				12-14-17	0.0057	13-14-15-18	0.0014
				12-15-16	0.0071	13-14-16-17	0.1048
				12-16-17	0.0028	13-14-16-18	0.0099
				12-16-18	0.0014	13-14-16-19	0.0014
				13-14-15	0.0496	13-14-17-18	0.0043
				13-14-16	0.0581	13-15-16-17	0.0283
				13-14-17	0.0198	13-15-16-18	0.0071
				13-14-18	0.0043	13-15-16-19	0.0014
				13-15-16	0.0255	14-15-16-17	0.0043
				13-15-17	0.0128		
h	0.9595			*h*	0.9668		

[a] One allele of these haplotypes has the deletion mutations.

[b] Only two copies of DYS464 are presented due to the deletion of the other two copies in one sample at 11-17 and in two samples at 14-16.

Table 10. Diversity, mutation rate and AMOVA analysis for each Y-STR marker

Y-STR	Gene diversity	Mutation rate ($x10^{-3}$)[a]	% Variance	
			Among haplogroups	Within haplogroups
DYS19	0.7115	1.76	44.82	55.18
DYS385	0.9595	2.09	15.00	85.00
DYS389-I	0.6675	2.44	46.07	53.93
DYS389-II	0.7244	2.60	19.04	80.96
DYS390	0.6733	2.29	44.43	55.57
DYS391	0.2667	3.11	24.87	75.13
DYS392	0.6789	0.68	76.17	23.83
DYS393	0.6281	0.81	66.17	33.83
DYS437	0.4320	2.02	65.59	34.41
DYS438	0.6330	0.33	84.84	15.16
DYS439	0.6329	5.37	12.80	87.20
DYS448	0.7510	0.00	61.27	38.73
DYS456	0.5108	5.59	31.34	68.66
DYS458	0.7779	8.38	10.91	89.09
DYS635	0.6857	5.66	26.24	73.76
GATA H4.1	0.6115	3.43	38.28	61.72
DYS388	0.4850	0.00	72.33	27.67
DYS446	0.7886	2.71	31.95	68.05
DYS447	0.7520	5.41	42.78	57.22
DYS449	0.8523	18.97	14.86	85.14
DYS464	0.9668	3.99	16.71	83.29

[a] Mutation rate information was reported by Lee et al.[49]

(B) Atypical and duplicated alleles at Y-STRs and their haplogroup memberships

Thirty two microvariant alleles were observed and their structures are described in Table 11. DYS449 contained microvariant alleles most frequently among 22 Y-STRs. According to the recent updated ISFG guidelines[41], it is

recommended that the allele should be designated by the intermediate number of the repeats when variations occur in integral core repeats. In the present study, the alleles, 27.2, 28.2, 29.2, 30.2 at DYS449; 14.1, 17.2 at DYS458; and 12.3, 14.3 at DYS464 were all caused from variations in integral numbers of repeats. Alleles 18 and 19 at DYS447 had a deletion of TAAAA-(TAATA) or (TAATA)-TAAAA within the core repeat unit. However, three variant alleles, resulting from mutations in flanking regions, were also found at the DYS449 loci. The three showed ostensible alleles of 25, 30.1 and 42, besides having genuine alleles, 26, 30, and 33, respectively. Sequence analysis revealed that alleles 25 and 42 had a 4 bp TCTC deletion and 36 bp duplication, respectively, at the flanking region located between the two core repeat units, DYS449.1 and DYS449.2. The variant allele 30.1 also had a thymidine insertion in the flanking region; it can be designated as DYS449.1*16 + DYS449.2*14 (U4Tins) according to guidelines mentioned above[41]. Duplications were observed once at each of DYS19, DYS390 and DYS447 loci (Figure 20) and were confirmed by cloning and sequencing. The sequence structures were shown in Table 11.

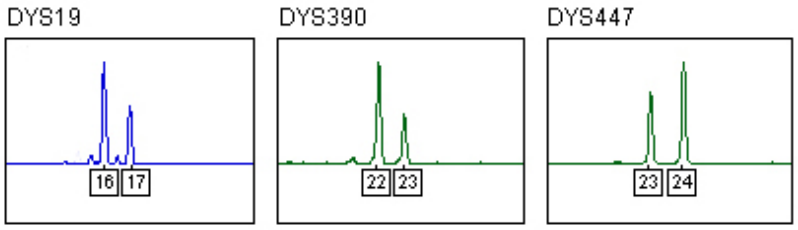


Figure 20. Electropherograms for duplicated alleles at the single copy loci DYS19, DYS390 and DYS447.

Table 11. Observed atypical and duplicated allele with their haplogroup membership

Y-STR	Allele[a]	Structure[b]	Haplogroup	N
Observed atypical alleles				
DYS447	18	$(TAATA)_7TAAAA(TAATA)_{10}$*$TAAAA(TAATA)_n$*	O3a1-KL2	2
	19	$(TAATA)_7TAAAA(TAATA)_{11}$*$TAAAA(TAATA)_n$*	O3a1-KL2	4
DYS449	25 (26)	$(TTTC)_{12}$-*tctc*tctcctcctc-$(tttc)_2$-N4-$(tttc)_2$-N12-cttc- $(TTTC)_{14}$	J1-M267	1
	27.2	$(TTTC)_3$**TC**$(TTTC)_{10}$-tctctctcctcctc-$(tttc)_2$-N4-$(tttc)_2$-N12-cttc-$(TTTC)_{14}$	O2b1-47z	1
	28.2	$(TTTC)_3$**TT**$(TTTC)_{11}$-tctctctcctcctc-$(tttc)_2$-N4-$(tttc)_2$-N12-cttc-$(TTTC)_{14}$	C3-M217	1
	29.2	$(TTTC)_3$**TC**$(TTTC)_{11}$-tctctctcctcctc-$(tttc)_2$-N4-$(tttc)_2$-N12-cttc-$(TTTC)_{15}$	O2b1-47z	1
	30.1 (30)	$(TTTC)_{16}$-tctctctcctcctc-$(tttc)_2$-N4-$(tttc)_2$-N12-c**t**ttc-$(TTTC)_{14}$	O3a1c-JST002611	1
	30.2	$(TTTC)_{16}$-tctctctcctcctc-$(tttc)_2$-N4-$(tttc)_2$-N12-cttc-$(TTTC)_{10}$**TT**$(TTTC)_4$	O1a1-P203	12
	42 (33)	$(TTTC)_{15}$-tctctctcctcctc-$[(tttc)_2$-N4-$(tttc)_2$-N12-cttc$]_2$-$(TTTC)_{18}$	O3*-M122	1
DYS458	14.1	$(GAAA)_{14}$**G**	N-M231	1
	17.2	$(GAAA)_{15}$**AA**$(GAAA)_2$	J1-M267	1
DYS464	12.3	$(CCTT)_7$**CTT**$(CCTT)_5$	O3a2*-P201	1
	14.3	$(CCTT)_3$**CTT**$(CCTT)_{11}$	N-M231	5
Observed duplicated alleles				
DYS19	16,17	$(TAGA)_3$tagg$(TAGA)_{13,14}$	O2b*-SRY_{465}	1
DYS390	22,23	$(TCTG)_8(TCTA)_{9,10}(TCTG)_1(TCTA)_4$	O3a2c1-M134	1
DYS447	23,24	$(TAATA)_7TAAAA(TAATA)_{7,8}TAAAA(TAATA)_7$	O3a2c1a- M117	1

[a] The numbers in parentheses are genuine alleles identified by sequence analysis.
[b] Deletion and insertion are indicated in *italic* and **bold**, respectively.

Haplogroup memberships were determined in samples with observed atypical and duplicated alleles (Table 11). The unusually short allele at DYS447, intermediate allele 30.2 at DYS449 and allele 14.3 at DYS464 were present in each single haplogroup background, haplogroups O3a1-KL2 (xO3a1c), O1a1-P203 and N-M231, respectively. Unusually short alleles at DYS447, which are caused by partial deletion of sequences $TAAAA(TAAATA)_n$, only occurred in haplogroup O3a1-KL2 (xO3a1c). At DYS449, the intermediate alleles caused by TC partial repeat insertion in the first TTTC track, named 27.2 and 29.2, were all detected in haplogroup O2b1-47z. Another type of intermediate allele is a TT partial repeat insertion in the first TTTC and the

second TTTC track, respectively. A sample with allele 28.2 from the first track insertion of TT belonged to haplogroup C3-M217 and all samples with allele 30.2 by the second track insertion belonged to haplogroup O1a1-P203. Finally, intermediate alleles, named 12.3 and 14.3 at DYS464 occurred in haplogroups O3a2*-P201 and N-M231, respectively.

Network analyses for Y-STR haplotypes carrying the observed atypical alleles were performed to evaluate the relationships among these haplotypes within a haplogroup context (Figure 21). The haplotypes excluding the Y-STRs that displayed atypical alleles were used in the network analysis to eliminate bias from that Y-STR. The haplotypes possessing the short allele variant at DYS447 formed a distinct cluster in the haplogroup O3a1-KL2 (xO3a1c) (Figure 21a), suggesting this variant likely defines a sublineage within the haplogroup O3a1-KL2 (xO3a1c). For DYS449 variants, only the haplotypes with allele 30.2 were closely related to every other variant, indicating shared ancestry (Figure 21b). Another network was constructed to approximate the relationship among the haplotypes with or without allele 30.2. The network was generated based on the analysis of Y-STR haplotypes belonging to haplogroup O1a1-P203 in the present study and a Japanese' haplotype with allele 30.2 obtained from the Sorenson Molecular Genealogy Foundation (SMGF) by manual searching (Figure 21c). The haplotypes carrying allele 30.2 formed a distinct cluster, indicating the possibility of defining a sublineage within haplogroup O1a1-P203. Unlike DYS447 and DYS449, the haplotypes with

allele 14.3 at DYS464 did not form a distinct branch (Figure 21d). This finding suggests that the variant allele 14.3 is partitioned across more than two sublineages within haplogroup N-M231.

Interestingly, an intermediate allele 17.2 at DYS458, which is known to be associated with haplogroup J1[27], was found in these samples. The corresponding haplogroup J1 was confirmed by an additional SBE reaction (Figure 17e).

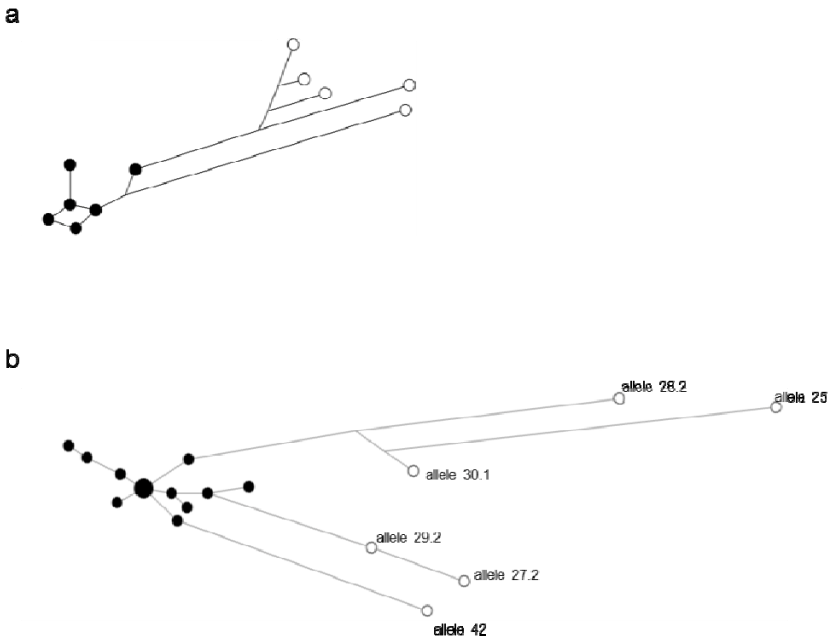


Figure 21. Median-network analysis of haplotypes carrying atypical alleles based on 17 Y-STR (DYS19, DYS388, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS446, DYS447, DYS448, DYS449, DYS456, DYS458, DYS635 and GATA H4.1) or 16 Y-STR loci except for the Y-STR displaying the atypical allele. Relationships between haplotypes belonging to haplogroup O3a1-KL2 (xO3a1c). (a) The haplotypes possessing the unusually short allele variant at DYS447 are indicated by filled circles. Relationships between haplotypes carrying DYS449 atypical alleles. (b)

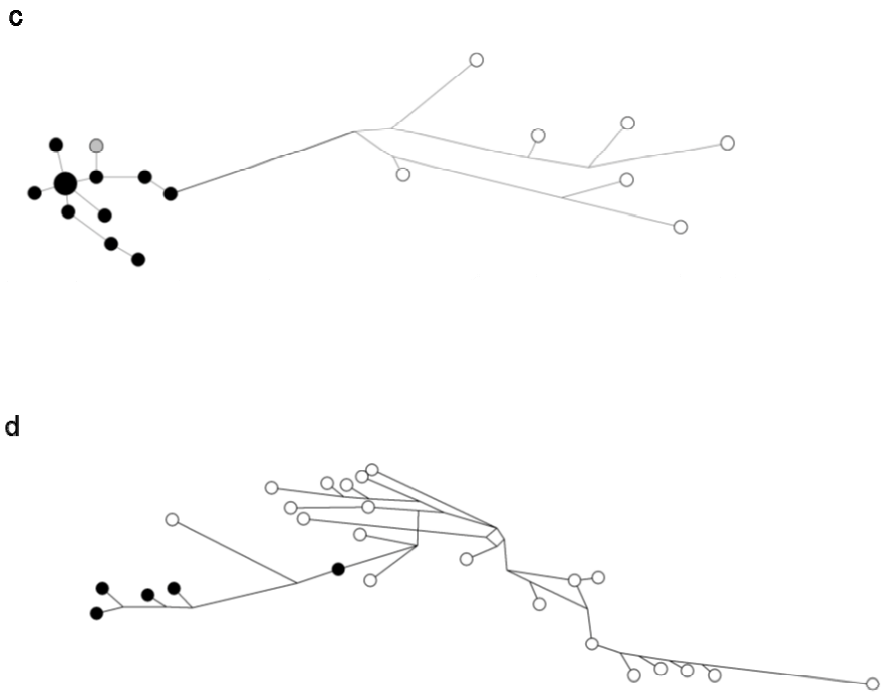


Figure 21. Cont. Relationships between haplotypes within haplogroup O1a1-P203. (c) The haplotypes possessing the allele 30.2 at DYS449 are indicated by filled circles. Relationships between haplotypes belonging to haplogroup N-M231. (d) The haplotypes possessing the allele 14.3 at DYS464 are indicated by filled circles.

(C) Deletions in the DYS385 flanking region and their haplogroup memberships

In the present study, DYS385 amplicon primers of multiplex STR-I/II and multiplex STR-III encompass different flanking regions from each other (Figure 22), and accordingly, they display different allelic designations when deletion or insertion mutations occur in one of these flanking regions. Allele designation repeat number discrepancies between the primer sets were observed in eight

(1.13%) of 706 individuals (Figure 23). Among these, three individuals showed a two-repeat unit difference and five individuals showed a one-repeat difference. However, PCR analysis using locus-discrimination primers for DYS385 revealed that these variations are not linked to a specific locus (DYS385a or DYS385b).

Sequence analysis of these eight samples revealed deletion mutations at two sites in the upstream flanking region of the DYS385 core repeat units, $(GAAA)_n$. One deletion was an 8 bp deletion of GAGAAAAA and the other was a 4 bp deletion of AAGG; these were found in blocks of $(GAGAAAAA)_2$ and $(AAGG)_6$, respectively (Figure 22). Therefore, primer set of multiplex STR-I/II resulted in two or one repeat differences from those of primer set of multiplex STR-III by amplifying the two deletion sites along with the DYS385 core repeat unit. The Powerplex® Y and AmpFlSTR® Yfiler™ kits produced the identical genotyping results as multiplex STR-I/II primer set of the present study. Without simultaneous use of two primer pairs, these flanking region variations could not be observed because the DYS385 core repeat unit is a tetramer.

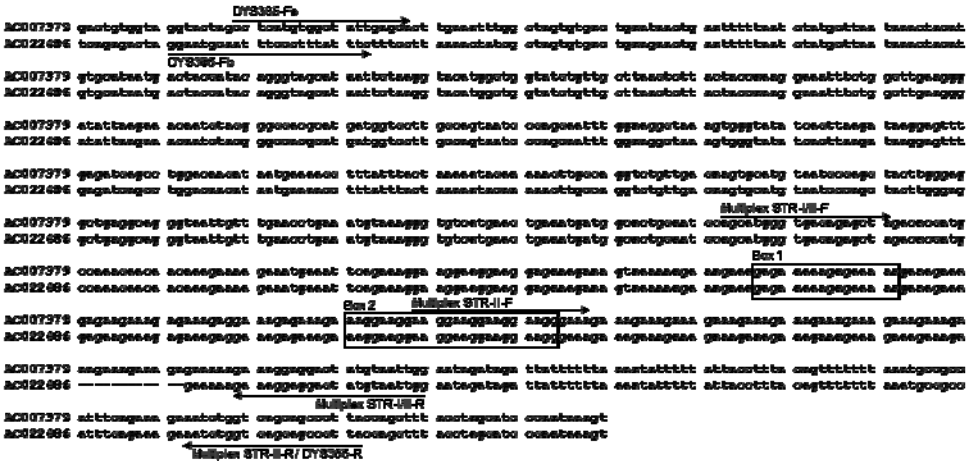


Figure 22. Sequence structure comparison of DYS385a (AC007379) and DYS385b (AC022486). Arrows indicate amplification primers used in the present study. Boxed sequences illustrate the two deletion sites. One is an 8 bp (GAGAAAAA) deletion and the other is a 4 bp (AAGG) deletion.

Y-chromosomal haplogroups were determined for eight samples with the deletion mutations in the DYS385 flanking region (Table 12). Among the two kinds of mutations at DYS385, the deletion of GAGAAAAA in the upstream of core repeat unit $(GAAA)_n$ only occurred within haplogroup O3a2b-M7. Network analysis of the haplotypes with the two deletion mutations revealed a distinct cluster formed by the haplotypes with an 8-bp deletion mutation (Figure 24). This suggests that the haplotypes with the 8-bp deletion have a common ancestry.

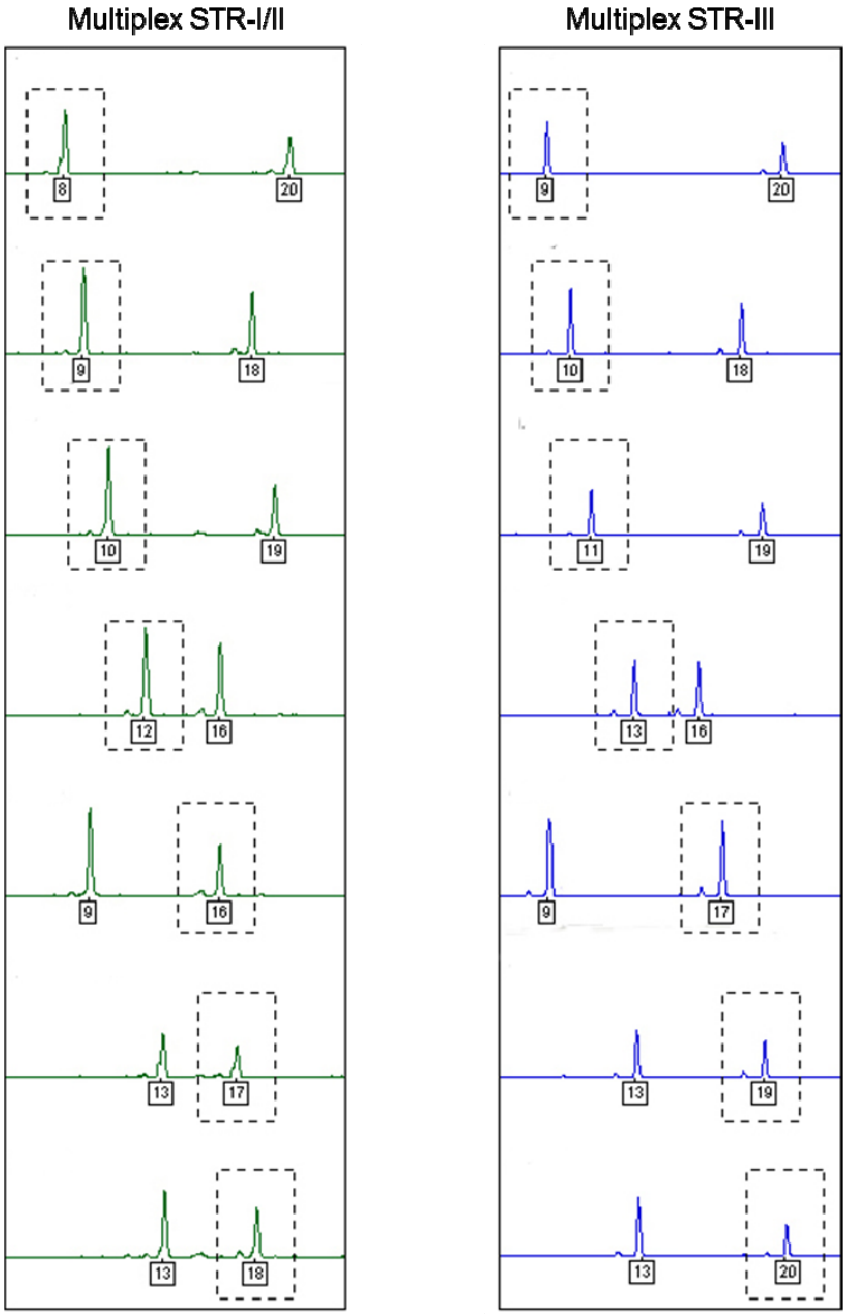


Figure 23. Electropherograms showing DYS385 allele designation discrepancies of one repeat or two repeats between multiplex STR-I/II and multiplex STR-III.

Table 12. Deletion mutations in DYS385 flanking region and their haplogroup membership

Allele[a]	Structure	Haplogroup	N
8(9)-20	$(gagaaaaa)_2$-N38-$(aagg)_5$-$(GAAA)_9$	O2b1-47z	1
9-16(17)	$(gagaaaaa)_2$-N38-$(aagg)_5$-$(GAAA)_{17}$	D-M174	1
9(10)-18	$(gagaaaaa)_2$-N38-$(aagg)_5$-$(GAAA)_{10}$	O2b*-SRY_{465}	1
10(11)-19	$(gagaaaaa)_2$-N38-$(aagg)_5$-$(GAAA)_{11}$	C3-M217	1
12(13)-16	$(gagaaaaa)_2$-N38-$(aagg)_5$-$(GAAA)_{13}$	D-M174	1
13-17(19)	$(gagaaaaa)_1$-N38-$(aagg)_6$-$(GAAA)_{19}$	O3a2b-M7	2
13-18(20)	$(gagaaaaa)_1$-N38-$(aagg)_6$-$(GAAA)_{20}$	O3a2b-M7	1

[a] The numbers in parentheses are genuine alleles identified by sequence analysis.

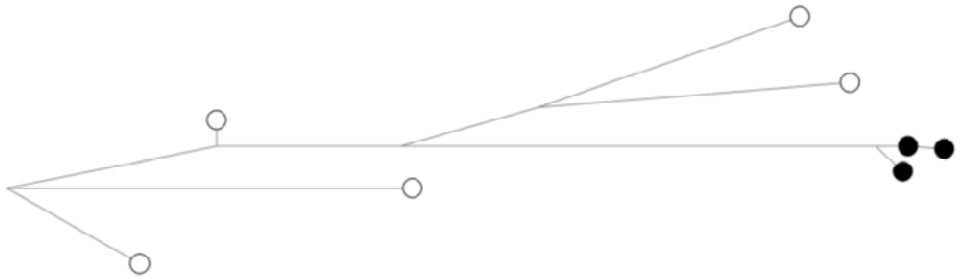


Figure 24. Relationship between the haplotypes carrying deletion mutations in the DYS385 flanking region. Median-network analysis of haplotypes carrying atypical alleles based on 17 Y-STR (DYS19, DYS388, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS446, DYS447, DYS448, DYS449, DYS456, DYS458, DYS635 and GATA H4.1). The haplotypes possessing the 8-bp deletion mutation are indicated by filled circles. The haplotypes possessing the 4-bp deletion mutation are indicated by open circles.

(D) Null alleles associated with *AZFc* microdeletions and their haplogroup memberships

Six samples showed a null allele at DYS448. Primers were designed for larger PCR fragments of DYS448, but failed to produce amplicons in the six samples. This supported that the null allele resulted from deletion of the entire

DYS448 locus. Accordingly, I further characterized the DYS448 null allele in Koreans using PCR assay for five STS markers around the DYS448 locus (Figure 25).

DYS448 is located within the azoospermia factor c gene (*AZFc)* in the distal euchromatic part of the Y chromosome long arm, Yq11.223.[50-52] *AZFc* consists almost entirely of very long direct and inverted repeats (Figure 25a). Therefore, it is prone to partial deletions or duplications by rearrangements.[52] A multi-locus STR known to be highly informative, DYS464, also lies within the r1-r4 ampliconic repeats of *AZFc* (Figure 25a). From a multiplex PCR assay for the six DYS448 null allele samples using the five STS marker set, which included sY1161, sY1191, sY1291, sY1206, and sY1201, three samples were found to have a b1/b3 deletion (Figure 25b, lanes 3-5) and the other three were found to have a polymorphic 50f2/C deletion (Figure 25b, lanes 6-8). Inferring from the alignment of direct and inverted repeats in *AZFc*, I expected that the three samples carrying a b1/b3 deletion would carry a simultaneous two-copy deletion at DYS464 (i.e. r1 and r2). Genotyping results confirmed this hypothesis, with two peaks of similar height at DYS464; one carried 11-17 haplotype and two carried 14-16 haplotype.

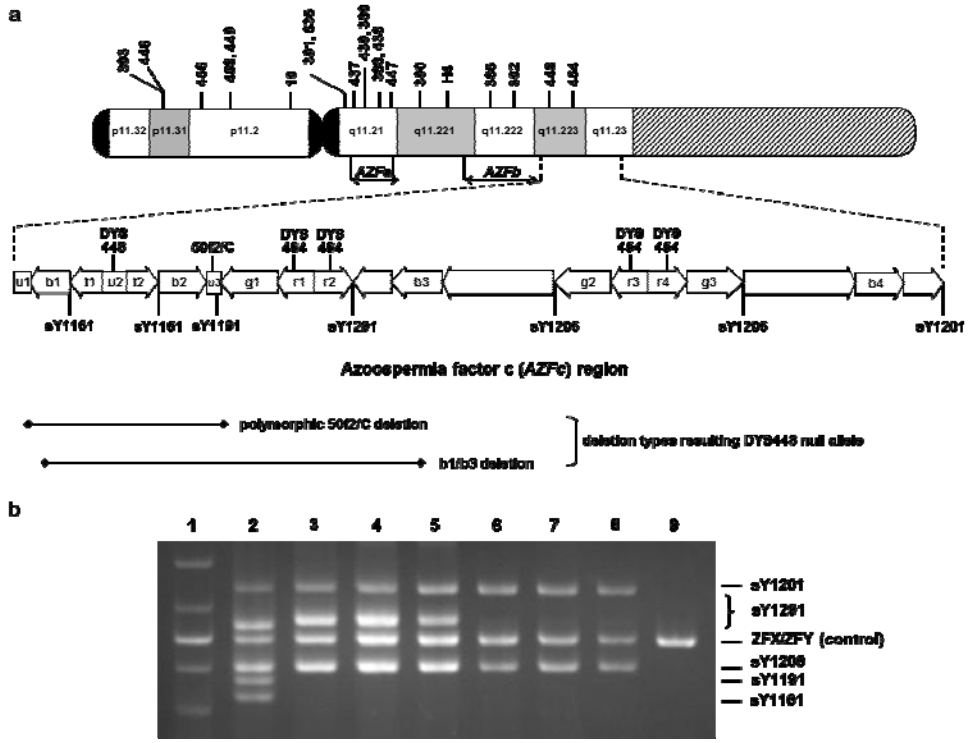


Figure 25. Y chromosome ideogram showing the approximate positions of *AZF*a, *AZF*b, *AZF*c, and the 22 Y-STR markers. The locations of DYS448, DYS464, and a series of STSs are indicated on the *AZF*c amplicon structure (a). Multiplex PCR assay of *AZF*c markers (b). Lane 1, 100 bp ladder size marker; Lane 2, a positive control without a null allele at DYS448; Lanes 3-8, six samples with a null allele at DYS448; Lane 9, a female control.

Y-chromosomal haplogroups were determined for six samples with the null alleles (Table 13). The null allele at DYS448 occurred in two different haplogroups, C3-M217 and O3a2*-P201. To assess the relationship of the haplotypes carrying the deletion by rearrangement, a network was constructed using both the haplotypes from this study and the reported haplotypes with the DYS448 deletion mutation[53] (Figure 26). Three haplotypes with a polymorphic

50f2/C deletion and the reported haplotypes with similar deletion mutation (called non-b1/b3 class II) all belonged to haplogroup C3-M217, although they formed two subgroups. Network results suggested the possibility of one common origin for this mutation. In contrast, haplotypes with the b1/b3 deletion mutation were wide spread in haplogroup C3-M217, O3a2*-P201 (in this study), C3c-M148, C*-RPS4Y$_{711}$ and O3a2c1-M134, indicating an independent origin of the deletion event. These results were consistent with those from the report by Balaresque et al.[53]

Table 13. Null alleles associated with *AZFc* microdeletion and their haplogroup membership

DYS448	DYS464	Type of deletion in *AZFc* region	Haplogroup	N
Null	11-12-17	polymorphic 50f2/C deletion	C3-M217	3
Null	14-16	b1/b3 deletion	O3a2*-P201	2
Null	11-17	b1/b3 deletion	C3-M217	1

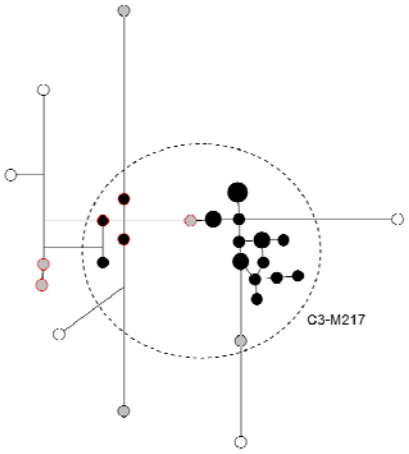


Figure 26. Relationship between haplotypes carrying DYS448 null alleles. The network was constructed using the haplotypes from the present study and Balaresque et al.[53] based on information from 11 Y-STR loci (DYS19, DYS388, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS446, DYS447, DYS449, DYS456, DYS458, DYS635 and GATA H4.1). The haplotypes with polymorphic 50f2/C deletion (or non-b1/b3 class II) are indicated by filled circles. The haplotypes with b1/b3 deletion are indicated by gray circles. The haplotypes with the other deletion type are indicated by open circles. The haplotypes from the present study are indicated by additional red circles.

C. Y-STR haplotypes and their relationship with haplogroups

In haplotype analysis for 22 Y-STRs, a total of 693 different haplotypes were identified among 706 unrelated Koreans. Of 693 haplotypes, 682 haplotypes (98.12%) were unique, 9 haplotypes were observed twice, and two haplotypes were observed three times (Table 14). Number of haplotypes, haplotype diversities and discriminatory capacities for the minimal haplotype, the extended SWGDAM haplotype, and the AmpFlSTR® Yfiler™ loci haplotype are shown in Table 15.

Table 14. Distributions of 22 Y-STRs haplotypes and their haplogroup membership in a Korean population

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H001	1	13	11,12	12	14	29	22	10	15	14	14	10	10	17	25	19	29	15	14.1	13,14	21	20	N
H002	1	13	12,16	12	12	28	23	10	12	12	15	11	12	13	23	19	33	15	19	13,14,16	19	21	O3a2c1*
H003	1	13	12,16	12	12	28	23	10	12	12	15	11	12	13	23	19	34	15	18	13,14,16	19	21	O3a2c1*
H004	1	13	13,14	13	13	30	23	10	11	13	14	10	14	13	24	17	34	15	20	14,16,18	21	20	C
H005	1	13	13,14	13	13	30	23	10	11	13	15	10	13	13	24	17	33	16	20	14,16,18	21	20	C
H006	1	13	14,22	12	14	30	23	9	14	13	14	12	11	14	27	20	30	15	16	15,16	22	19	Q
H007	1	13	14,22	12	14	30	23	9	14	13	14	12	11	14	27	20	30	15	16	15,16	23	19	Q
H008	1	13	15,20	12	14	29	23	9	15	13	14	12	11	14	27	19	30	16	18	14,15	22	19	Q
H009	1	13	15,20	12	14	29	23	9	15	14	14	12	11	14	27	19	30	16	18	14,15	22	19	Q
H010	1	13	15,20	12	14	30	24	9	14	14	14	12	11	14	27	19	30	16	17	15,16	22	19	Q
H011	1	13	15,20	12	14	30	26	9	15	14	14	12	12	14	26	19	30	16	17	15	22	19	Q
H012	1	13	15,21	12	14	31	24	9	14	14	14	12	11	15	27	19	31	15	18	15,16	23	19	Q
H013	1	13	15,22	12	13	30	24	9	14	14	14	12	11	14	26	19	31	16	16	15,16	22	19	Q
H014	1	13	15,22	12	14	31	24	9	14	14	14	12	11	15	27	19	31	16	16	15,16	23	19	Q
H015	1	13	15,23	12	14	30	23	9	14	14	14	12	11	14	25	19	29	15	16	16	22	19	Q
H016	1	13	16,20	12	14	29	24	9	15	14	14	12	11	14	27	19	30	16	18	14,15	22	19	Q
H017	1	13	19,19	12	14	30	24	9	14	14	14	12	11	15	27	19	30	15	16	15,16	22	19	Q
H018	1	14	10,18	12	13	28	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16,17	21	20	O2b*
H019	1	14	10,18	12	14	28	23	10	13	13	14	13	13	14	25	18	29	15	17	13,14,17	20	20	O2b*
H020	1	14	10,18	12	14	29	24	10	13	13	14	13	11	12	25	19	30	15	18	13,14,16,17	21	20	O2b*
H021	1	14	10,19	12	14	30	22	10	14	13	14	13	12	14	25	18	32	15	19	13,15,16,17	20	21	O2b1
H022	1	14	10,21	12	14	30	22	10	13	13	14	13	12	14	24	19	30	15	18	14,15,16,17	21	22	O2b1
H023	1	14	10,21	12	15	31	22	10	13	13	14	13	12	14	24	19	30	15	18	14,15,16,17	21	22	O2b1
H024	1	14	11,12	12	14	31	22	10	14	13	14	10	10	14	25	20	31	15	17	13	21	21	N
H025	1	14	11,12	13	13	29	23	11	14	13	14	10	11	13	26	20	28	17	17	12,13,14	24	21	N
H026	1	14	11,12	13	14	30	23	10	14	13	14	10	11	14	26	20	28	18	16	12	24	21	N
H027	1	14	11,12	13	14	30	23	11	14	13	14	10	11	14	26	20	28	18	18	12,13,14	23	21	N
H028	1	14	11,12	13	14	30	23	11	15	13	14	10	14	13	25	20	26	17	16	13	22	19	N
H029	1	14	11,12	13	14	31	23	10	14	13	14	10	11	14	26	19	28	17	16	12,13,14	23	21	N
H030	1	14	11,12	13	14	31	23	10	14	13	14	10	12	14	26	19	28	17	16	12,13,14	23	21	N
H031	1	14	11,12	13	14	32	23	10	14	13	14	10	11	15	26	20	29	18	17	12,13	24	21	N
H032	1	14	11,13	12	13	29	23	11	14	14	14	10	10	15	26	19	27	14	18	14,14.3	21	21	N
H033	1	14	11,13	12	14	30	23	11	14	13	14	10	10	18	26	18	30	14	18	15,16	23	20	N
H034	1	14	11,13	12	14	30	23	11	16	14	14	11	10	17	25	19	28	14	16	14,14.3	21	21	N
H035	1	14	11,13	12	14	30	23	11	16	14	14	11	11	18	25	19	28	15	17	14,14.3	21	21	N
H036	1	14	11,13	12	14	31	22	10	13	13	15	10	11	15	27	19	29	14	18	13	21	21	N
H037	1	14	11,13	12	14	31	23	11	16	14	14	11	10	17	25	20	27	14	17	14,14.3	22	21	N
H038	1	14	11,13	13	14	31	23	10	14	12	14	10	11	14	26	20	28	15	15	12,13	22	21	N
H039	1	14	11,13	13	14	31	23	10	14	13	14	10	11	14	25	20	28	18	17	12,13,14	23	21	N
H040	1	14	11,14	12	13	30	21	10	15	13	14	10	11	14	26	19	29	16	17	13	21	21	N
H041	1	14	11,14	12	13	30	21	10	15	13	14	10	11	14	26	19	30	14	17	13	21	21	N
H042	1	14	11,14	12	14	30	23	11	16	14	14	11	11	19	25	19	29	14	17	14,14.3	21	21	N
H043	1	14	11,14	13	13	29	21	10	15	13	14	10	10	16	26	19	30	16	17	13	21	21	N
H044	1	14	11,17	12	13	28	23	10	11	14	14	10	11	17	25	21	29	15	18	13,14,16	23	20	C3
H045	1	14	11,17	12	13	28	24	10	14	13	14	10	11	16	25	19	29	15	17	10,14,16	20	20	O2b*

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H046	1	14	12,12	12	13	29	24	11	16	13	14	10	13	15	24	19	31	15	16	13,14,16,17	21	21	N
H047	1	14	12,13	12	13	29	23	10	15	13	14	10	11	14	28	19	28	15	16	13,15	21	21	N
H048	1	14	12,16	12	12	30	24	10	12	12	15	10	11	13	24	19	33	15	19	13,15	19	21	O3a2c1*
H049	1	14	12,18	10	12	28	25	10	15	12	15	11	13	13	23	20	32	15	18	12,13,15	20	20	O3a2c1a*
H050	1	14	12,18	10	12	29	25	10	15	12	15	11	13	13	23	20	31	15	18	12,13,15	20	20	O3a2c1a*
H051	1	14	12,19	10	12	27	24	10	14	12	15	11	11	14	23	20	32	15	19	12,13,14,15	20	21	O3a2c1a*
H052	1	14	12,19	10	13	29	23	10	14	12	15	11	12	13	23	20	31	15	17	12,13,15	20	22	O3a2c1a*
H053	1	14	12,19	10	14	30	24	10	14	12	15	11	12	13	23	19	31	15	18	12,13,15	20	21	O3a2c1a*
H054	1	14	13,13	10	12	28	24	10	14	12	15	11	12	13	24	20	32	15	17	11,13,15	21	21	O3a2c1a*
H055	1	14	13,14	15	14	31	24	9	11	12	14	10	14	13	24	20	25	14	17.2	13,15,16,18	21	19	J1
H056	1	14	13,16	10	12	27	24	10	13	12	15	11	12	13	23	20	32	16	15	12,14,15	20	21	O3a2c1a*
H057	1	14	13,16	10	12	27	24	10	14	12	15	11	11	13	23	20	34	15	16	12,13,14,15	20	21	O3a2c1a*
H058	1	14	13,17	10	12	27	24	10	14	12	14	11	12	13	23	20	33	15	18	13,15	21	21	O3a2c1a*
H059	1	14	13,17	10	12	28	25	10	14	12	15	11	11	13	23	20	32	14	18	12,13,15	20	21	O3a2c1a*
H060	1	14	13,17	10	12	28	25	10	14	12	15	11	12	13	23	20	33	15	17	12,13,14	20	20	O3a2c1a*
H061	1	14	13,17	10	13	29	25	11	14	12	15	11	12	13	23	18	35	15	17	12,14,15	20	22	O3a2c1a*
H062	1	14	13,17	10	14	29	25	11	14	12	15	11	12	13	23	18	34	15	17	12,14,15	20	22	O3a2c1a*
H063	1	14	13,18	10	12	26	24	10	14	12	15	11	12	13	23	20	32	15	17	13,14,15	21	21	O3a2c1a*
H064	1	14	13,18	10	12	27	24	10	14	12	14	11	11	13	23	20	32	15	17	12,13,14,15	20	21	O3a2c1a*
H065	1	14	13,18	10	12	27	24	10	14	12	14	11	12	13	23	20	34	15	17	13,15	22	22	O3a2c1a*
H066	1	14	13,18	10	12	27	24	10	14	12	15	11	11	13	23	19	32	15	19	12,13,14,15	21	22	O3a2c1a*
H067	1	14	13,18	10	12	27	24	10	14	12	15	11	11	13	24	20	32	15	19	12,13,15	21	21	O3a2c1a*
H068	1	14	13,18	10	12	27	24	10	14	12	15	11	13	13	23	20	29	15	17	13,15	20	21	O3a2c1a*
H069	1	14	13,18	10	12	27	24	10	14	12	15	11	13	13	24	20	31	15	18	13	21	22	O3a2c1a*
H070	1	14	13,18	10	12	27	24	10	14	13	15	11	11	14	23	20	34	15	17	12,14,15	21	21	O3a2c1a*
H071	1	14	13,18	10	12	28	23	10	14	12	15	11	12	13	23	19	33	16	16	12,14,15	22	22	O3a2c1a*
H072	1	14	13,18	10	12	28	24	10	14	12	15	11	12	12	23	20	31	13	17	12,14,15	20	21	O3a2c1a*
H073	1	14	13,18	10	12	28	24	10	14	12	15	11	12	14	23	20	34	15	18	12,13,14,16	20	20	O3a2c1a*
H074	1	14	13,18	10	12	28	24	10	14	13	14	11	11	14	24	21	33	13	17	12,14,15	20	21	O3a2c1a*
H075	1	14	13,18	10	12	28	24	11	14	12	15	11	12	12	23	20	30	14	17	12,13,14,16	20	21	O3a2c1a*
H076	1	14	13,18	10	12	28	24	11	14	12	16	11	13	13	24	20	33	15	16	12,14,16	20	21	O3a2c1a*
H077	1	14	13,18	10	12	28	25	10	13	12	15	11	13	14	23	20	33	15	19	12,13,14,15	20	22	O3a2c1a*
H078	1	14	13,18	10	12	28	25	10	15	12	15	11	14	13	23	20	32	15	18	11,12,13,15	20	20	O3a2c1a*
H079	1	14	13,18	10	12	29	24	10	14	12	15	11	13	14	25	20	32	15	19	13,15	21	21	O3a2c1a*
H080	1	14	13,18	10	13	28	23	10	14	12	15	11	13	14	23	20	32	15	17	13,14,16	20	20	O3a2c1a*
H081	1	14	13,18	10	14	30	25	11	14	12	15	11	13	13	23	18	34	15	18	12,14,15	20	22	O3a2c1a*
H082	1	14	13,18	12	13	30	25	11	13	13	15	11	12	16	25	19	42	15	16	13,14,15	21	20	O3*
H083	1	14	13,18	12	14	30	24	10	13	13	14	10	9	15	24	20	30	15	17	15,16	21	21	O3*
H084	1	14	13,19	10	11	28	25	10	14	12	15	11	12	13	23	20	34	15	19	12,13,15	20	21	O3a2c1a*
H085	1	14	13,19	10	11	28	25	10	14	12	15	11	12	13	23,24	20	34	15	20	12,13,15	20	21	O3a2c1a*
H086	1	14	13,19	10	12	27	23	10	14	12	15	11	12	13	23	20	32	14	19	12,13,14,15	20	21	O3a2c1a*
H087	1	14	13,19	10	12	27	24	10	10	12	15	11	12	13	23	20	34	14	17	13,14	20	20	O3a2c1a*
H088	1	14	13,19	10	12	27	24	10	10	13	15	11	12	13	24	20	33	15	17	13,15	21	21	O3a2c1a*
H089	1	14	13,19	10	12	27	24	10	14	12	15	11	11	13	23	20	34	15	18	12,13,14,15	20	21	O3a2c1a*
H090	1	14	13,19	10	12	28	23	10	14	12	15	11	13	14	23	19	32	16	15	12,14,15	21	22	O3a2c1a*
H091	1	14	13,19	10	12	28	24	10	14	12	15	10	12	13	23	20	34	15	18	12,13,15	20	21	O3a2c1a*
H092	1	14	13,19	10	12	28	24	10	14	12	15	11	12	12	23	20	31	15	19	12,13,14,15	20	21	O3a2c1a*

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H093	1	14	13,19	10	12	28	24	10	14	12	15	11	12	13	23	20	35	16	18	12,13,15	20	21	O3a2c1a*
H094	1	14	13,19	10	12	28	24	10	14	12	15	11	12	13	23	21	33	15	19	12,13,15	20	21	O3a2c1a*
H095	1	14	13,19	10	12	28	24	10	14	12	15	11	12	14	23	20	32	15	19	13,14,15	22	22	O3a2c1a*
H096	1	14	13,19	10	12	28	24	10	14	12	16	11	13	13	23	20	32	15	16	12,14,15	20	21	O3a2c1a*
H097	1	14	13,19	10	12	28	25	10	14	12	14	11	12	13	23	20	31	15	20	12,13,15	20	21	O3a2c1a*
H098	1	14	13,19	10	12	28	25	10	14	12	15	11	11	13	24	20	33	15	21	12,13,15	20	21	O3a2c1a*
H099	1	14	13,19	10	12	28	25	10	14	12	15	11	12	14	23	20	34	15	18	12,13,15	20	21	O3a2c1a*
H100	1	14	13,19	10	12	28	25	10	14	12	15	11	12	16	23	20	33	15	18	12,13,15	20	21	O3a2c1a*
H101	1	14	13,19	10	12	28	25	10	14	12	15	11	13	13	23	20	34	15	18	12,13,15	20	22	O3a2c1a*
H102	1	14	13,19	10	12	29	25	10	14	12	15	11	12	13	23	20	32	15	18	12,13,15	20	21	O3a2c1a*
H103	1	14	13,19	10	12	30	25	10	14	12	15	11	12	14	23	20	32	15	18	12,13,15	20	21	O3a2c1a*
H104	1	14	13,19	10	14	29	24	10	14	12	15	11	11	13	23	20	34	15	18	12,13,14,15	20	21	O3a2c1a*
H105	1	14	13,20	10	12	27	23	10	14	12	15	11	11	14	23	20	33	15	19	12,13,14,15	20	21	O3a2c1a*
H106	1	14	13,20	10	12	28	23	10	14	12	15	11	11	13	24	20	33	15	17	13,14,15	20	21	O3a2c1a*
H107	1	14	13,20	10	12	28	24	10	14	12	14	11	12	15	23	20	32	15	19	13,14,15	22	22	O3a2c1a*
H108	1	14	13,20	10	12	28	24	10	14	12	15	11	11	13	23	20	34	15	17	12,13,14,15	21	21	O3a2c1a*
H109	1	14	13,20	10	12	28	24	10	14	12	15	11	11	13	24	20	34	15	17	12,13,14,15	21	21	O3a2c1a*
H110	1	14	13,20	10	12	28	24	10	14	12	15	11	12	13	23	20	34	15	18	12,13,14	20	20	O3a2c1a*
H111	1	14	13,20	10	12	28	24	10	14	12	15	11	12	14	23	20	32	15	18	13,14,15	22	22	O3a2c1a*
H112	1	14	13,20	10	12	28	24	10	14	12	15	11	12	14	23	20	33	15	17	13,14,15	22	22	O3a2c1a*
H113	1	14	13,20	10	12	28	24	10	15	12	15	11	12	13	24	20	31	15	19	12,13,15	20	21	O3a2c1a*
H114	1	14	13,20	10	12	28	25	11	14	12	15	11	11	13	23	20	32	15	18	13,15	20	21	O3a2c1a*
H115	1	14	13,21	10	12	28	24	10	14	12	15	11	12	13	23	20	32	15	18	13,14,15	21	21	O3a2c1a*
H116	1	14	14,17	10	13	29	25	10	14	12	15	11	11	12	23	19	32	15	18	12,13,14,15	20	21	O3a2c1a*
H117	1	14	14,18	10	12	28	24	11	15	13	15	11	12	14	25	20	36	14	17	13,14,15	20	21	O3a2c1a*
H118	1	14	14,18	10	14	30	25	11	14	12	15	11	12	13	23	18	34	15	17	12,14,15	20	22	O3a2c1a*
H119	1	14	14,19	10	11	27	23	10	14	12	15	11	11	13	23	22	32	15	17	12,13,14	21	21	O3a2c1a*
H120	1	14	14,19	10	11	27	23	10	14	12	15	11	12	13	24	21	31	15	18	12,13,14	20	21	O3a2c1a*
H121	1	14	14,19	10	12	29	24	10	14	12	15	12	12	13	23	20	33	17	18	13,14,15	21	21	O3a2c1a*
H122	1	14	14,20	10	11	27	23	10	14	12	15	11	13	13	23	20	32	15	19	12,13,14	20	20	O3a2c1a*
H123	1	14	14,20	10	12	28	24	11	14	12	14	11	12	13	24	20	34	15	17	12,13,15	20	21	O3a2c1a*
H124	1	14	15,19	10	11	27	23	10	14	12	15	11	12	13	23	21	33	15	18	12,13,14	20	21	O3a2c1a*
H125	1	14	15,20	10	12	28	23	10	14	12	15	11	12	11	23	20	32	15	16	12,13,15	21	21	O3a2c1a*
H126	1	14	15,21	12	13	29	24	9	14	14	14	12	12	14	27	19	31	16	16	15	22	19	Q
H127	1	15	8,20	14	13	30	22	10	13	13	14	13	11	14	26	18	30	15	18	13,15,16	20	21	O2b1
H128	1	15	9,18	12	14	31	22	10	13	13	14	13	12	14	25	18	30	15	20	12,15,16,17	20	21	O2b1
H129	1	15	9,19	12	13	28	22	10	13	13	14	13	12	13	25	18	30	15	19	13,15,16	20	21	O2b1
H130	1	15	9,19	12	13	29	22	10	13	13	14	13	11	14	24	18	29	15	18	13,15,16,19	20	21	O2b1
H131	1	15	9,19	12	13	29	22	10	13	13	14	13	12	15	25	17	31	15	19	13,15,16,17	20	21	O2b1
H132	1	15	9,19	12	14	30	22	10	13	13	14	13	12	14	25	17	31	15	19	13,16	20	21	O2b1
H133	1	15	9,19	12	14	30	22	10	13	13	14	13	12	16	25	18	31	15	18	13,15,16	20	21	O2b1
H134	1	15	9,19	12	14	30	22	10	13	13	14	14	12	14	25	18	30	15	18	13,16	20	21	O2b1
H135	1	15	9,19	12	14	31	22	10	13	13	14	13	12	14	25	18	30	15	19	13,15,16	20	21	O2b1
H136	1	15	9,19	12	14	31	22	10	13	13	14	13	12	14	25	18	30	15	20	13,15,16	20	21	O2b1
H137	1	15	9,19	12	14	31	22	10	13	13	14	13	13	14	25	17	32	15	17	13,15,16	20	21	O2b1
H138	1	15	9,20	12	13	29	22	10	13	13	14	13	12	13	25	18	30	15	20	13,15,16	20	21	O2b1
H139	1	15	10,17	12	13	28	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16,17	21	21	O2b*

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H140	1	15	10,17	12	13	28	23	10	14	13	14	13	12	13	25	18	30	15	17	13,14,16,17	21	20	O2b*
H141	1	15	10,17	12	13	29	23	10	13	13	14	13	12	13	25	18	30	15	18	13,14,16	21	20	O2b*
H142	1	15	10,17	12	14	28	23	10	13	13	14	13	13	14	25	18	28	15	18	13,16,18	20	21	O2b*
H143	1	15	10,17	12	14	29	23	10	13	13	14	13	12	14	25	18	33	15	17	13,16	24	20	O2b*
H144	1	15	10,17	12	14	29	23	10	14	13	14	13	11	14	25	18	33	15	18	13,14,16,17	21	20	O2b*
H145	1	15	10,17	12	14	29	24	10	13	13	14	13	12	12	25	18	30	15	19	13,14,16	21	20	O2b*
H146	1	15	10,17	12	14	30	22	10	13	13	14	13	12	14	25	18	29	15	18	13,16,17	21	21	O2b1
H147	1	15	10,17	12	14	30	22	10	13	13	14	13	12	14	25	18	30	15	19	13,16,17	20	21	O2b1
H148	1	15	10,17	12	14	30	22	10	13	13	14	13	12	14	25	19	30	15	19	13,15,16,17	20	21	O2b1
H149	1	15	10,17	12	14	30	22	10	13	13	14	13	12	14	26	18	30	16	19	13,15,16,17	21	21	O2b1
H150	1	15	10,18	12	13	27	23	10	13	13	14	13	12	14	24	18	28	15	17	13,16	20	21	O2b*
H151	1	15	10,18	12	13	27	23	10	13	13	14	13	12	14	24	18	28	15	17	13,14,16	20	21	O2b*
H152	1	15	10,18	12	13	27	23	10	13	13	14	13	13	14	25	18	28	15	16	13,16,18	20	21	O2b*
H153	1	15	10,18	12	13	27	23	10	13	13	14	13	13	14	25	18	30	15	16	13,16,17	20	22	O2b*
H154	1	15	10,18	12	13	28	22	10	13	13	14	13	12	14	25	18	30	14	19	13,15,17	20	21	O2b1
H155	1	15	10,18	12	13	28	23	10	13	13	14	13	13	14	25	18	30	15	17	13,16,18	20	21	O2b*
H156	1	15	10,18	12	13	29	23	10	13	13	14	13	12	13	26	18	31	15	18	13,16,17	21	20	O2b*
H157	1	15	10,18	12	13	30	22	10	13	14	14	13	12	15	25	18	30	15	19	13,16,18	20	21	O2b*
H158	1	15	10,18	12	14	28	23	10	13	13	14	13	13	14	25	18	28	16	16	13,16,17	20	21	O2b*
H159	1	15	10,18	12	14	28	23	10	13	13	14	13	13	14	25	18	29	15	16	13,16,18	21	21	O2b*
H160	1	15	10,18	12	14	29	22	10	13	13	14	13	11	14	25	18	30	16	17	13,14,16,17	20	20	O2b*
H161	1	15	10,18	12	14	29	22	10	13	13	14	13	12	14	25	18	30	15	17	13,15,16	20	20	O2b1
H162	1	15	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	32	15	17	13,14,16,17	21	20	O2b*
H163	1	15	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	33	15	18	13,14,16,17	21	20	O2b*
H164	1	15	10,18	12	14	29	23	10	13	13	14	13	13	13	25	18	31	15	17	14,16,17	21	20	O2b*
H165	1	15	10,18	12	14	29	23	11	13	13	14	13	12	12	25	18	30	15	18	13,14,16,17	21	20	O2b*
H166	1	15	10,18	12	14	29	24	10	13	13	14	13	13	13	25	19	30	15	17	13,14,16,17	21	20	O2b*
H167	1	15	10,18	12	14	29	24	10	13	14	14	11	12	14	25	18	31	15	18	13,14,16,18	21	20	O2b*
H168	1	15	10,18	12	14	30	22	10	13	13	14	13	11	14	25	18	28	15	18	12,15,16,18	21	21	O2b1
H169	1	15	10,18	12	14	30	22	10	13	13	14	13	12	13	25	18	29	15	20	13,16	20	21	O2b1
H170	1	15	10,18	12	14	30	22	10	13	13	14	13	12	14	24	18	29	15	19	13,15,16,18	20	21	O2b1
H171	1	15	10,18	12	14	30	22	10	13	13	14	13	12	14	24	18	30	15	19	13,15,16,18	20	22	O2b1
H172	1	15	10,18	12	14	30	22	10	13	13	14	13	12	14	24	18	30	16	18	13,15,16,18	20	21	O2b1
H173	1	15	10,18	12	14	30	22	10	13	13	14	13	12	14	25	18	29	15	18	16,17	20	21	O2b1
H174	1	15	10,18	12	14	30	22	10	13	13	14	13	12	14	25	18	30	15	19	13,15,16,17	20	21	O2b1
H175	1	15	10,18	12	14	30	22	10	13	13	14	13	12	15	25	18	29	15	19	13,15,16,17	21	21	O2b1
H176	1	15	10,18	12	14	30	22	10	13	13	14	13	12	15	25	18	30	15	19	13,16,17	20	21	O2b1
H177	1	15	10,18	12	14	30	22	10	13	13	14	13	12	16	25	18	29	15	18	13,16,17	20	21	O2b1
H178	1	15	10,18	12	14	30	22	10	13	13	14	13	13	15	25	18	30	15	17	13,15,16,17	20	21	O2b1
H179	1	15	10,18	12	14	31	22	10	13	13	14	13	11	14	25	18	30	15	17	13,16,17	20	21	O2b1
H180	1	15	10,18	12	14	32	22	10	13	13	15	13	12	14	25	18	29	15	18	13,16,17	21	21	O2b1
H181	1	15	10,18	13	13	27	23	10	13	13	14	13	11	14	25	18	28	18	17	13,16,18	20	21	O2b*
H182	1	15	10,18	13	13	27	23	10	13	13	14	13	13	15	25	18	28	15	16	13,16,17	20	21	O2b*
H183	1	15	10,18	13	15	30	23	10	13	13	14	13	10	13	25	18	30	15	18	13,14,16,17	21	20	O2b*
H184	1	15	10,19	12	13	27	23	11	13	13	14	13	14	14	25	18	27	15	18	13,16,18	21	21	O2b*
H185	1	15	10,19	12	13	28	22	10	13	14	14	13	12	14	26	17	31	13	18	13,15,16,17	21	21	O2b1
H186	1	15	10,19	12	13	28	23	10	13	13	14	13	12	14	25	19	30	15	17	13,15,17	20	21	O2b1

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H187	1	15	10,19	12	13	28	24	10	13	13	14	13	12	13	25	18	32	15	17	13,14,16,17	21	20	O2b*
H188	1	15	10,19	12	13	29	22	10	13	13	14	13	12	14	26	18	29	15	18	13,15,16,17	19	21	O2b1
H189	1	15	10,19	12	13	29	22	10	13	13	15	13	13	14	25	18	31	15	17	13,15,16,17	21	21	O2b1
H190	1	15	10,19	12	14	28	23	10	13	13	14	13	12	14	25	18	28	15	15	13,16,17	20	21	O2b*
H191	1	15	10,19	12	14	29	22	11	14	13	14	13	13	14	25	18	30	15	19	13,15,16,17	20	21	O2b1
H192	1	15	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16,17	21	20	O2b*
H193	1	15	10,19	12	14	29	23	11	13	13	14	13	12	13	25	18	30	15	18	13,14,16,17	21	20	O2b*
H194	1	15	10,19	12	14	29	24	10	11	15	14	10	12	16	27	21	32	15	16	12,14,16	21	20	C3
H195	1	15	10,19	12	14	29	24	10	11	15	14	10	12	16	27	21	33	15	15	12,14,16	21	20	C3
H196	1	15	10,19	12	14	29	24	10	13	13	14	13	12	13	25	18	32	15	19	15,17	21	20	O2b*
H197	1	15	10,19	12	14	30	22	10	13	13	14	13	11	15	25	18	29	15	19	13,15,16,17	20	21	O2b1
H198	1	15	10,19	12	14	30	22	10	13	13	14	13	12	14	25	18	30	14	18	14,15,16,17	20	21	O2b1
H199	1	15	10,19	12	14	30	22	10	13	13	14	13	12	14	26	18	30	15	20	13,15,17	21	20	O2b1
H200	1	15	10,19	12	14	30	22	10	13	13	14	13	12	15	25	18	30	15	19	13,15,16	20	21	O2b1
H201	1	15	10,19	12	14	30	22	10	13	13	14	13	12	16	25	18	30	15	20	16,17	20	20	O2b1
H202	1	15	10,19	12	14	30	23	10	11	14	14	10	12	19	26	21	30	16	18	12,14,15	21	20	C3
H203	1	15	10,19	12	14	30	23	10	13	13	14	13	12	16	25	18	30	15	20	16,17	20	20	O2b1
H204	1	15	10,19	12	14	31	22	10	13	13	14	13	12	14	25	18	30	15	17	12,15,16,17	20	21	O2b1
H205	1	15	10,19	12	14	31	23	10	13	13	14	13	12	13	25	18	29	16	19	13,17,18	20	21	O2b1
H206	1	15	10,19	12	14	31	23	10	13	14	14	14	13	13	25	18	29	16	19	13,17,18	21	21	O2b1
H207	1	15	10,19	12	15	31	22	10	13	14	14	13	12	14	25	17	30	16	17	13,15,16	21	21	O2b1
H208	1	15	10,19	13	14	30	22	10	13	13	14	13	13	14	25	18	29	15	19	12,15,16,17	21	21	O2b1
H209	1	15	10,20	11	14	30	23	10	13	13	14	13	11	16	24	17	27.2	15	17	12,15,17,18	20	22	O2b1
H210	1	15	10,20	12	13	29	22	10	13	13	14	13	11	13	25	18	30	15	18	13,15,16,17	20	21	O2b1
H211	1	15	10,20	12	13	29	22	10	13	13	14	13	11	14	25	18	30	14	18	13,15,16	20	21	O2b1
H212	1	15	10,20	12	14	29	22	10	13	13	14	13	11	13	25	18	29	15	19	13,16,17	20	22	O2b1
H213	1	15	10,20	12	14	29	24	10	11	15	14	10	11	16	27	20	31	15	15	12,14,16	22	20	C3
H214	1	15	10,20	12	14	30	22	10	13	13	14	12	12	14	23	18	29	15	20	13,15,16,17	20	21	O2b1
H215	1	15	10,20	12	14	30	22	10	13	13	14	13	12	14	25	18	30	15	17	13,14,16,17	20	21	O2b1
H216	1	15	10,20	12	14	30	22	10	13	13	14	13	12	15	25	18	30	15	17	13,15,16	20	21	O2b1
H217	1	15	10,20	12	14	30	22	10	14	13	14	13	12	14	27	18	32	15	18	13,15,17	20	21	O2b1
H218	1	15	10,20	12	14	30	22	10	14	13	14	13	12	14	27	18	32	15	18	13,15,18	20	20	O2b1
H219	1	15	10,20	12	14	30	23	10	13	13	14	13	11	13	24	17	29.2	14	17	13,15,17	20	21	O2b1
H220	1	15	10,20	12	14	30	23	10	13	13	14	13	12	14	25	18	31	15	18	13,16,17	20	21	O2b1
H221	1	15	10,20	12	14	30	23	10	13	13	14	14	12	13	24	18	32	15	17	13,14,16,17	22	20	O2b*
H222	1	15	10,20	12	14	32	22	10	13	13	14	13	12	14	25	18	29	15	19	13,14,16,18	21	21	O2b1
H223	1	15	10,20	12	15	31	22	10	13	13	14	13	11	15	25	18	29	15	18	13,16,17	21	22	O2b1
H224	1	15	10,21	12	13	28	22	10	13	13	14	13	12	13	26	17	31	13	17	13,15,16,17	20	21	O2b1
H225	1	15	10,21	12	14	30	22	10	13	13	14	13	12	14	25	18	30	15	19	13,15,16,17	21	21	O2b1
H226	1	15	11,11	12	13	28	22	10	14	12	15	10	12	11	26	18	31	15	17	13,15	20	21	O3a1c
H227	1	15	11,12	12	14	29	23	10	14	13	14	10	11	13	25	18	30	14	15	14,16	20	20	N
H228	1	15	11,12	12	14	30	21	10	14	13	14	10	11	15	25	18	29	14	18	15,17	20	21	N
H229	1	15	11,12	12	15	30	23	10	11	14	14	10	12	16	26	21	29	15	17	12,14,15,18	21	20	C3
H230	1	15	11,15	12	14	29	23	10	11	14	14	10	12	17	24	21	30	15	17	12,16,17	24	20	C3
H231	1	15	11,16	12	13	28	23	10	12	12	14	10	12	13	24	19	34	15	21	12,13,14,15	19	21	O3a2c1*
H232	1	15	11,16	13	13	29	23	10	11	15	14	10	13	15	26	21	29	14	13	12,15,16	21	20	C3
H233	1	15	11,17	12	13	29	23	10	11	14	15	10	12	17	25	21	29	15	18	13,17	23	20	C3

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H234	1	15	11,17	12	14	29	23	10	11	14	14	10	12	17	25	21	29	15	18	13,16	23	20	C3
H235	1	15	11,17	12	14	29	23	10	11	14	14	10	12	17	26	20	28	15	16	12,13,16	23	20	C3
H236	1	15	11,17	12	14	29	23	10	11	14	14	10	12	17	26	21	29	15	17	12,13,16,17	22	20	C3
H237	1	15	11,17	12	14	29	23	10	11	14	14	10	12	17	26	21	30	15	17	12,13,16,18	20	20	C3
H238	1	15	11,17	12	14	29	23	10	11	14	14	10	13	17	26	21	29	15	17	12,13,15,17	21	20	C3
H239	1	15	11,17	12	14	29	23	10	11	14	14	11	13	17	25	21	29	14	17	13,17	22	20	C3
H240	1	15	11,17	13	13	30	23	10	11	13	14	10	12	16	27	20	30	15	17	12,13,14,16	21	20	C3
H241	1	15	11,18	12	12	29	23	10	12	12	15	10	11	13	24	19	33	16	18	13,14,16	21	20	O3a2c1*
H242	1	15	11,18	12	13	30	23	10	12	12	14	10	12	13	24	19	34	15	18	12,13,15	19	21	O3a2c1*
H243	1	15	11,18	12	14	30	22	10	13	13	14	13	11	14	24	18	30	16	18	13,15,18	20	21	O2b1
H244	1	15	11,18	12	14	30	23	10	11	15	14	10	12	18	26	21	30	15	16	12,13,16,17	21	20	C3
H245	1	15	11,18	13	14	30	23	10	11	15	14	10	10	15	25	21	30	15	15	12,13,14,17	21	20	C3
H246	1	15	11,18	13	14	31	23	10	11	13	14	10	12	16	25	21	30	15	16	12,13,14,17	21	20	C3
H247	1	15	11,19	12	12	28	23	10	12	12	16	10	11	13	24	19	32	16	18	13,14,16	21	20	O3a2c1*
H248	1	15	11,19	12	12	29	23	10	12	12	15	10	11	13	24	19	32	16	17	13,14,16	21	20	O3a2c1*
H249	1	15	11,19	12	12	29	23	10	12	12	15	10	11	13	24	19	32	16	18	13,14	21	20	O3a2c1*
H250	3	15	11,19	12	12	29	23	10	12	12	15	10	11	13	24	19	32	16	18	13,14,16	21	20	O3a2c1*
H251	1	15	11,19	12	12	29	23	10	12	12	15	10	11	13	24	19	32	16	18	13,14,16	22	20	O3a2c1*
H252	1	15	11,19	12	12	29	23	10	12	12	15	10	11	13	24	19	32	16	18	13,14,15,16	21	20	O3a2c1*
H253	1	15	11,19	12	12	29	23	10	12	12	15	10	11	13	24	19	32	16	19	13,14,16	21	20	O3a2c1*
H254	2	15	11,19	12	12	29	23	10	12	12	15	10	11	13	24	19	33	16	18	13,14,16	21	20	O3a2c1*
H255	1	15	11,19	12	12	29	23	10	12	12	15	10	11	13	24	19	33	16	19	13,14,16	21	20	O3a2c1*
H256	1	15	11,19	12	12	29	23	10	12	12	15	10	11	14	24	19	32	16	18	13,14,16	21	20	O3a2c1*
H257	1	15	11,19	12	12	29	23	10	12	12	15	10	12	13	24	19	32	16	18	13,14,16	22	20	O3a2c1*
H258	1	15	11,19	13	13	29	20	11	11	16	14	10	11	17	26	21	29	15	16	12,14,15	19	20	C3
H259	1	15	11,19	13	13	29	23	10	11	15	14	10	10	15	25	21	30	15	15	13,14	21	20	C3
H260	1	15	11,19	13	13	29	25	10	11	14	14	10	10	18	25	20	30	15	16	12,13,14,15	23	21	C3
H261	1	15	11,19	13	13	29	25	10	11	14	14	10	10	18	26	20	30	15	16	12,13,14,15	23	21	C3
H262	1	15	11,19	13	14	30	23	10	11	15	14	10	10	15	25	21	30	15	15	12,14	21	20	C3
H263	1	15	11,19	13	14	30	23	10	12	15	14	10	10	15	25	21	30	15	15	12,13,14,17	21	20	C3
H264	1	15	11,19	13	14	30	23	11	11	15	14	10	10	15	25	21	31	15	15	12,13,14,17	21	20	C3
H265	1	15	11,20	12	12	29	23	10	12	12	15	10	11	13	24	19	32	16	18	14,16	21	20	O3a2c1*
H266	1	15	11,20	12	12	29	23	10	12	12	15	10	12	13	24	19	32	16	19	13,14,15	22	20	O3a2c1*
H267	1	15	11,20	12	12	29	23	10	12	12	15	10	12	13	24	19	35	16	19	13,14,16	21	20	O3a2c1*
H268	1	15	11,20	13	13	29	23	10	11	14	14	10	10	15	25	21	30	16	15	12,13,14	21	20	C3
H269	2	15	11,20	13	14	30	23	10	11	15	14	10	10	15	25	21	30	15	15	12,13,14	21	20	C3
H270	1	15	11,21	13	14	30	23	10	11	14	14	10	13	14	24	21	32	15	16	12,13,14,16	21	20	C3
H271	1	15	12,12	12	12	29	23	10	12	12	15	10	11	14	24	19	34	15	18	13,15	19	21	O3a2c1*
H272	1	15	12,13	12	12	28	23	11	14	13	14	10	12	13	25	18	30.2	12	17	13,14	20	21	O1a1
H273	1	15	12,14	12	12	27	25	9	14	14	14	10	12	12	24	20	30	15	15	13,14	21	20	O1a1
H274	1	15	12,14	12	12	28	24	10	14	14	14	10	12	12	24	19	29	16	15	13,14	22	20	O1a1
H275	1	15	12,14	12	12	29	23	10	13	12	15	10	12	12	24	20	34	14	17	13,15	20	21	O3a2c1*
H276	1	15	12,14	13	13	29	24	10	11	13	14	11	12	15	25	0	29	17	16	11,12,17	20	20	C3
H277	1	15	12,14	13	13	29	24	11	11	13	14	11	12	15	26	0	29	17	16	11,12,17	20	20	C3
H278	1	15	12,14	13	13	30	24	9	11	13	14	10	12	14	29	21	31	15	17	11,12,13,16	22	20	C3
H279	1	15	12,15	12	12	29	23	10	12	12	15	10	13	12	23	19	32	16	17	13,15	19	21	O3a2c1*
H280	1	15	12,15	12	12	29	24	10	14	14	14	10	12	13	24	18	30	16	16	13,14	21	20	O1a1

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H281	1	15	12,16	12	11	27	23	10	12	12	15	10	12	13	25	19	33	15	17	12,13,15	20	21	O3a2c1*
H282	1	15	12,16	12	12	28	23	10	12	12	14	10	11	11	25	19	33	15	17	13,14,15	19	20	O3a2c1*
H283	2	15	12,16	12	12	28	23	10	12	12	15	10	12	13	24	19	34	15	18	13,15	20	22	O3a2c1*
H284	1	15	12,16	12	12	28	23	10	12	12	15	10	12	13	25	19	33	15	18	13,15	20	22	O3a2c1*
H285	1	15	12,16	12	12	29	23	10	12	12	15	10	11	13	24	19	33	15	19	12,13,15	20	22	O3a2c1*
H286	1	15	12,16	12	12	29	23	11	12	12	15	10	13	14	24	19	31	15	17	13,14,15	20	21	O3a2c1*
H287	1	15	12,16	12	13	29	23	10	13	12	15	10	13	13	24	19	32	15	17	12,13,14,15	20	21	O3a2c1*
H288	1	15	12,16	12	13	30	24	10	13	14	14	10	11	14	27	18	29	15	15	14,16	21	21	O2*
H289	1	15	12,16	12	14	30	24	9	13	13	14	10	11	14	26	18	29	15	15	14,16	21	20	O2*
H290	1	15	12,16	12	14	30	25	10	13	14	14	10	13	14	28	18	28	15	14	14,15,16	21	21	O2*
H291	1	15	12,16	12	14	31	24	10	13	13	14	10	11	14	26	17	29	15	15	14,16	22	20	O2*
H292	1	15	12,16	13	13	29	24	10	11	13	14	11	13	15	26	0	29	17	16	11,12,17	22	20	C3
H293	1	15	12,17	12	12	28	23	10	12	12	14	10	12	13	24	19	31	15	18	13,14,16	19	21	O3a2c1*
H294	1	15	12,17	12	12	28	23	10	12	12	15	10	11	13	24	19	30	14	19	12,13,15	21	21	O3a2c1*
H295	1	15	12,17	12	12	28	23	10	12	13	14	10	12	12	25	19	33	17	17	13,15	19	22	O3a2c1*
H296	1	15	12,17	12	12	28	24	10	13	12	14	10	12	12	25	19	32	14	18	13,14,15	22	21	O3a1c
H297	2	15	12,17	12	12	29	23	10	12	13	14	10	11	12	25	19	34	17	16	13,15	19	22	O3a2c1*
H298	2	15	12,17	12	12	29	23	10	12	13	14	10	11	12	25	19	34	17	17	13,15	19	22	O3a2c1*
H299	1	15	12,17	12	12	29	23	10	12	13	14	10	12	12	25	19	33	16	17	13,15	19	22	O3a2c1*
H300	1	15	12,17	12	12	29	2223	10	13	12	15	10	12	14	24	19	33	15	19	13,15,16	20	21	O3a2c1*
H301	1	15	12,17	12	12	30	23	10	12	12	14	10	14	13	24	19	35	15	18	12,13,15	19	20	O3a2c1*
H302	1	15	12,17	12	13	29	24	10	13	13	14	10	11	14	27	18	28	15	15	14,15,16	21	21	O2*
H303	1	15	12,17	12	14	29	24	10	13	15	14	10	12	14	27	18	29	15	15	14,16	21	21	O2*
H304	1	15	12,17	12	14	31	24	10	13	13	14	10	11	14	27	18	29	15	15	14,15,17	21	21	O2*
H305	1	15	12,18	12	12	28	23	10	12	12	15	10	12	12	24	19	32	15	15	13,14,15	19	21	O3a2c1*
H306	1	15	12,18	12	12	28	24	10	13	12	14	10	11	13	24	19	32	15	18	13,14,16,17	23	21	O3a1c
H307	1	15	12,18	12	12	28	25	11	13	12	14	10	12	13	24	20	31	14	19	12,13,16,18	21	20	O3a1c
H308	1	15	12,18	12	12	29	23	10	12	13	14	10	12	12	25	19	33	18	17	13,15	19	22	O3a2c1*
H309	1	15	12,18	12	12	29	24	10	13	12	15	9	10	14	24	19	34	16	15	12,13,15	22	22	O3a2c*
H310	1	15	12,18	12	13	29	24	11	13	13	15	10	11	15	23	20	35	17	17	13,14,15	21	20	O3a2c*
H311	1	15	12,18	12	14	30	23	10	13	12	15	10	12	12	18	21	31	14	18	15,16	24	20	O3a1
H312	1	15	12,18	12	14	30	23	10	13	12	15	10	12	12	18	21	32	14	19	16	24	20	O3a1
H313	1	15	12,18	12	14	30	23	10	13	12	15	10	12	12	19	21	32	14	17	16	23	20	O3a1
H314	1	15	12,18	13	12	28	25	10	13	12	14	10	12	13	24	19	33	14	17	13,14,16,18	21	21	O3a1c
H315	1	15	12,18	13	14	30	23	10	11	15	14	10	11	16	25	21	34	16	15	12,16	21	20	C3
H316	1	15	12,19	10	12	27	23	10	14	13	15	11	12	13	23	20	32	15	19	12,13,15	20	21	O3a2c1a*
H317	1	15	12,19	10	12	28	25	10	14	12	15	11	12	12	23	20	29	15	15	12,13,14,16	20	21	O3a2c1a*
H318	1	15	12,19	10	13	29	24	10	14	12	15	11	12	13	24	20	33	15	18	12,13,15,16	20	20	O3a2c1a*
H319	1	15	12,19	12	12	28	24	10	13	12	15	9	10	13	23	19	31	15	16	12,13,14,15	21	20	O3a2c*
H320	1	15	12,19	12	12	29	23	10	12	13	15	10	12	12	24	19	32	15	19	13,14,15	19	21	O3a2c1*
H321	1	15	12,19	12	12	30	24	11	13	12	15	9	10	15	24	20	32	15	16	12,13,14,15	21	21	O3a2c*
H322	1	15	12,19	12	13	29	24	11	14	13	14	10	11	15	26	19	31	16	18	13,14,16	20	20	O2b*
H323	1	15	12,19	12	13	30	24	11	14	13	14	10	11	15	26	19	31	15	17	13,14,16,17	20	20	O2b*
H324	1	15	12,19	12	13	30	25	10	13	12	15	9	10	14	23	19	32	16	15	12,13,15	21	22	O3a2c*
H325	1	15	12,19	12	14	30	23	10	13	12	15	10	12	12	19	21	31	14	17	16	24	20	O3a1
H326	1	15	12,19	12	14	30	23	10	13	12	15	10	13	12	19	21	31	14	17	15,16	24	20	O3a1
H327	1	15	12,19	12	14	30	23	10	13	12	15	10	13	12	19	21	31	14	18	15,16	24	20	O3a1

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H328	1	15	12,19	12	14	31	23	10	12	12	15	10	12	11	25	19	33	15	18	13,14	19	21	O3a2c1*
H329	1	15	12,20	12	12	28	25	11	13	12	14	10	12	12	24	19	32	14	19	12,14,16,18	23	20	O3a1c
H330	1	15	12,20	13	13	29	23	10	11	15	14	10	10	15	25	21	30	15	15	12,13,14	22	20	C3
H331	1	15	13,13	12	12	28	23	10	14	13	14	10	11	13	25	18	30.2	12	17	13,14	20	21	O1a1
H332	1	15	13,13	12	12	28	23	11	14	13	14	10	11	13	25	18	30.2	12	17	13,13	20	21	O1a1
H333	2	15	13,13	12	12	28	23	11	14	13	14	10	12	13	25	18	30.2	12	17	13,14	20	21	O1a1
H334	1	15	13,13	12	12	28	23	11	14	13	14	10	12	13	25	18	30.2	12	18	13,14	21	21	O1a1
H335	1	15	13,13	12	12	28	23	11	14	13	14	10	12	14	25	18	30.2	12	17	13,13	20	21	O1a1
H336	1	15	13,13	12	12	28	23	11	14	13	14	11	12	13	25	18	30.2	12	16	13,14	20	21	O1a1
H337	1	15	13,13	12	12	28	23	11	16	13	14	10	11	14	25	18	30.2	12	17	13,14	20	21	O1a1
H338	1	15	13,13	12	12	28	23	11	16	13	14	10	12	14	25	18	30.2	12	17	13,14	20	21	O1a1
H339	1	15	13,13	12	12	29	23	11	14	13	14	10	11	12	25	18	31	17	15	13,14	19	20	O1a1
H340	1	15	13,14	12	12	29	23	11	14	13	14	10	12	13	26	18	30.2	12	17	13,14	20	21	O1a1
H341	1	15	13,14	13	13	29	24	10	11	13	14	11	12	15	29	0	25	18	17	11,17	21	20	C3
H342	1	15	13,16	12	12	29	22	11	12	12	15	10	12	14	24	19	33	16	20	13,15	20	21	O3a2c1*
H343	1	15	13,16	12	12	29	23	11	12	12	15	10	12	13	26	18	34	15	18	13,14	21	21	O3a2c1*
H344	1	15	13,16	12	13	28	22	10	13	14	14	10	11	15	26	21	28	17	16	13,15	23	20	O2*
H345	1	15	13,16	12	13	30	25	10	11	13	14	10	12	19	25	18	27	16	16	16,17	21	20	D
H346	1	15	13,17	12	12	28	23	10	12	13	15	10	12	12	25	19	31	14	18	13,14,15	19	21	O3a2c1*
H347	1	15	13,17	12	12	29	23	11	12	12	15	10	12	12	24	19	33	17	19	13,15	20	21	O3a2c1*
H348	1	15	13,17	12	12	29	24	10	13	12	15	10	13	13	24	19	35	15	18	13,17	19	21	O3a2c1*
H349	1	15	13,17	12	12	29	24	10	13	12	15	10	13	13	24	19	35	15	19	13,17	19	21	O3a2c1*
H350	1	15	13,17	12	12	29	25	10	13	12	14	10	12	14	24	20	27	14	17	12,13,15,17	24	21	O3a1c
H351	1	15	13,17	12	13	32	24	10	13	13	14	10	11	14	27	20	30	14	15	14,16	21	21	O2*
H352	1	15	13,18	10	12	28	25	10	14	12	15	11	12	13	24	20	30	15	17	12,13,14,15	20	22	O3a2c1a*
H353	1	15	13,18	12	12	28	23	10	12	13	15	10	12	11	25	19	32	15	17	13,14,15	19	21	O3a2c1*
H354	1	15	13,18	12	12	28	23	10	12	13	15	10	12	12	23	19	33	15	17	13,14,15	19	21	O3a2c1*
H355	1	15	13,18	12	12	29	23	10	12	13	15	10	12	12	24	19	30	15	18	13,14	19	21	O3a2c1*
H356	1	15	13,18	12	12	29	23	10	12	13	15	10	12	12	24	19	31	14	18	13,14,15	19	21	O3a2c1*
H357	1	15	13,18	12	12	29	23	10	12	13	15	10	12	12	24	19	33	15	17	13,14,15	19	21	O3a2c1*
H358	1	15	13,18	12	12	29	23	10	12	13	15	10	13	13	24	19	32	15	19	13,14,15	19	21	O3a2c1*
H359	1	15	13,18	12	12	29	23	11	12	12	15	10	12	13	24	19	36	16	17	13,15	21	21	O3a2c1*
H360	1	15	13,18	12	12	29	25	10	13	12	15	10	13	11	23	20	35	15	17	13,14,15	20	22	O3a2b
H361	1	15	13,18	12	13	30	23	11	12	12	15	10	13	13	24	19	33	15	17	13,15	20	21	O3a2c1*
H362	1	15	13,19	10	12	27	23	10	14	12	15	11	12	13	23	20	32	15	19	12,13,15	20	21	O3a2c1a*
H363	1	15	13,19	10	12	27	24	10	14	12	15	11	11	13	23	19	32	15	19	12,13,14,15	20	21	O3a2c1a*
H364	1	15	13,20	10	12	27	23	10	14	12	15	11	12	13	23	20	32	15	18	12,13,15	20	21	O3a2c1a*
H365	1	15	13,20	10	13	28	23	10	14	12	15	11	12	13	23	20	33	15	18	12,13,15	20	22	O3a2c1a*
H366	1	15	13,20	10	13	29	24	11	15	12	14	11	13	12	23	21	30	16	21	10,13,15	20	21	O3a2c1a*
H367	1	15	14,14	12	12	29	26	10	11	13	14	11	12	16	25	17	32	17	15	16,17	21	19	D
H368	1	15	14,16	12	12	29	23	10	12	12	15	10	12	13	23	19	34	17	18	13,14,15	21	21	O3a2c1*
H369	1	15	14,16	12	12	29	23	11	12	12	15	10	13	13	24	19	34	16	17	13,15,16	21	21	O3a2c1*
H370	1	15	14,17	12	12	27	24	10	13	12	14	10	11	13	24	20	29	13	18	13,14	23	21	O3a1c
H371	1	15	14,17	12	12	27	25	11	13	12	14	10	12	12	24	19	26	14	18	13,15,18	22	21	O3a1c
H372	1	15	14,17	12	12	29	25	10	13	12	14	10	11	14	24	20	28	14	17	12,13,15,17	24	21	O3a1c
H373	1	15	14,18	12	12	29	23	10	12	13	15	10	12	12	24	19	32	15	18	13,14,15	19	21	O3a2c1*
H374	1	15	14,18	12	12	29	25	10	13	12	14	10	11	15	24	20	28	14	18	13,14,17	24	21	O3a1c

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H375	1	15	14,18	13	12	30	24	10	13	12	15	10	10	11	24	20	32	15	17	12,14,15	21	22	O3a2*
H376	1	15	14,19	12	12	29	23	10	13	13	16	10	11	15	24	19	31	15	18	12,13,14,15	21	20	O3a2c*
H377	1	15	14,20	12	12	28	23	10	13	13	15	10	12	15	24	19	30	15	16	12,13,14,15	21	20	O3a2c*
H378	1	15	14,20	15	13	29	23	10	11	16	14	10	10	19	25	21	29	15	17	13,14,16	21	20	C3
H379	1	15	14,21	14	14	30	23	10	11	16	14	10	11	17	25	21	28	15	17	13,14,16	21	20	C3
H380	1	15	14,22	13	12	30	24	10	13	12	14	10	12	11	24	20	31	15	17	12,13,15,16	21	21	O3a2*
H381	1	15	15,15	12	12	27	24	11	13	13	15	10	13	13	23	19	31	15	19	12,13,15,16	22	21	O3a2c*
H382	1	15	15,17	12	13	30	25	10	11	13	14	10	12	16	25	17	33	16	15	17	21	21	D
H383	1	15	15,19	13	12	30	24	10	13	12	15	10	11	11	24	19	32	15	17	12,14,15,16	21	21	O3a2*
H384	1	15	15,20	13	12	29	24	10	13	12	15	10	13	11	24	20	31	15	16	12,14,15,16	21	21	O3a2*
H385	1	15	15,21	13	12	29	24	10	13	12	15	10	13	11	23	20	31	15	16	12,14,15,16	20	21	O3a2*
H386	1	15	15,21	13	12	30	24	10	13	12	15	10	11	13	24	20	31	15	17	12,13,16,19	21	21	O3a2*
H387	1	15	16,17	12	12	29	22	11	12	12	15	10	12	14	24	19	34	16	19	12,13,15,16	20	21	O3a2c1*
H388	1	16	9,18	12	14	29	24	10	14	13	14	13	13	14	25	18	32	15	17	13,14,16,17	21	20	O2b*
H389	1	16	9,19	12	14	30	24	11	13	13	14	12	12	14	24	18	31	15	18	13,16	20	21	O2b*
H390	1	16	9,19	14	14	30	24	11	13	13	15	12	12	15	24	18	29	14	17	12,16	20	21	O2b*
H391	1	16	9,20	12	14	30	23	11	13	13	14	12	12	14	24	18	29	15	17	13,16	21	21	O2b*
H392	1	16	10,10	12	14	29	23	11	12	13	14	13	12	13	25	18	30	15	17	13,14,16,17	22	20	O2b*
H393	1	16	10,16	12	13	28	23	10	13	13	14	13	11	13	25	18	29	15	17	14,15,16	22	20	O2b*
H394	1	16	10,16	12	14	29	23	10	13	13	14	13	12	13	25	19	30	15	19	12,16,17	21	20	O2b*
H395	1	16	10,16	13	14	29	23	10	13	13	15	13	12	13	26	18	30	15	17	12,16,18	21	20	O2b*
H396	1	16	10,17	12	13	28	23	10	13	13	14	13	12	13	25	18	30	15	17	13,14,17	21	20	O2b*
H397	1	16	10,17	12	13	28	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,15,17	21	20	O2b*
H398	1	16	10,17	12	13	28	23	10	13	13	14	13	12	13	27	18	30	15	18	13,14,16,18	21	21	O2b*
H399	1	16	10,17	12	13	28	23	10	13	13	14	13	13	13	25	18	31	15	18	13,14,16,17	21	20	O2b*
H400	1	16	10,17	12	13	28	24	10	13	13	14	13	12	13	25	18	30	14	17	13,14,16,17	22	20	O2b*
H401	1	16	10,17	12	14	28	23	10	12	13	14	13	11	12	24	18	30	15	17	13,14,16,17	22	20	O2b*
H402	1	16	10,17	12	14	29	23	10	13	13	14	13	10	13	25	18	30	15	17	13,14,16,17	21	20	O2b*
H403	1	16	10,17	12	14	29	23	10	13	13	14	13	11	13	25	18	32	16	16	13,14,17	21	20	O2b*
H404	1	16	10,17	12	14	29	23	10	13	13	14	13	11	13	25	18	33	15	17	13,14,16,17	22	19	O2b*
H405	1	16	10,17	12	14	29	23	10	13	13	14	13	11	13	26	18	30	15	17	13,14,16,17	21	21	O2b*
H406	1	16	10,17	12	14	29	23	10	13	13	14	13	12	13	25	18	30	15	18	13,14,17,18	21	20	O2b*
H407	1	16	10,17	12	14	29	23	10	13	13	14	13	12	13	25	18	33	15	16	14,16,17	21	20	O2b*
H408	1	16	10,17	12	14	29	23	10	13	13	14	13	12	14	25	18	27	15	17	13,14,17	24	20	O2b*
H409	1	16	10,17	12	14	29	23	10	13	13	14	13	13	12	25	18	30	15	17	13,14,17	22	20	O2b*
H410	1	16	10,17	12	14	29	23	10	13	13	14	13	13	14	25	18	31	15	17	13,14,16	21	20	O2b*
H411	1	16	10,17	12	14	29	24	10	13	13	14	13	12	12	25	18	31	15	17	13,14,17	21	20	O2b*
H412	1	16	10,17	12	14	30	23	10	13	10	14	14	12	13	25	18	30	15	18	14,16,17	21	20	O2b*
H413	1	16	10,17	12	14	30	23	10	13	13	14	13	12	13	25	18	30	15	19	13,14,16,17	21	20	O2b*
H414	1	16	10,17	12	14	30	23	10	13	13	14	14	11	12	25	18	33	15	18	13,14,16,17	21	20	O2b*
H415	1	16	10,17	13	13	28	23	10	13	13	14	13	12	13	25	18	32	15	17	12,13,16,18	21	20	O2b*
H416	1	16	10,18	12	13	28	22	10	13	13	14	13	12	13	25	18	29	15	15	13,14,16,17	21	20	O2b*
H417	1	16	10,18	12	13	28	23	10	13	13	14	13	12	12	26	18	30	15	18	13,14,16	22	20	O2b*
H418	1	16	10,18	12	13	28	23	10	13	13	14	13	12	13	25	18	30	16	17	13,14,16,17	21	20	O2b*
H419	1	16	10,18	12	13	28	23	10	13	13	14	13	12	13	25	18	31	15	16	13,14,16,17	20	20	O2b*
H420	1	16	10,18	12	13	28	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16,17	21	20	O2b*
H421	1	16	10,18	12	13	28	23	10	13	13	14	13	12	14	25	18	30	15	17	13,14,16,17	21	20	O2b*

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H422	1	16	10,18	12	13	28	24	10	13	13	14	13	12	13	26	18	30	15	17	13,14,16,17	22	20	O2b*
H423	1	16	10,18	12	14	28	23	10	13	13	14	13	12	13	25	18	30	15	18	13,14,17	21	20	O2b*
H424	1	16	10,18	12	14	29	23	9	13	13	14	13	12	13	26	18	31	15	19	13,14,16	21	20	O2b*
H425	1	16	10,18	12	14	29	23	10	13	13	14	13	9	13	25	18	30	16	18	13,14	21	20	O2b*
H426	1	16	10,18	12	14	29	23	10	13	13	14	13	11	13	25	18	31	15	18	13,14,16,17	21	19	O2b*
H427	1	16	10,18	12	14	29	23	10	13	13	14	13	11	13	25	18	32	15	17	13,14,16,17	21	20	O2b*
H428	1	16	10,18	12	14	29	23	10	13	13	14	13	11	14	25	19	31	15	18	13,14,16,18	21	20	O2b*
H429	1	16	10,18	12	14	29	23	10	13	13	14	13	12	12	25	18	30	15	18	13,14,16	22	20	O2b*
H430	1	16	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	30	15	17	13,14,16,17	21	20	O2b*
H431	1	16	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	30	15	18	13,14,16,17	22	20	O2b*
H432	1	16	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	18	12,14,16,17	21	20	O2b*
H433	1	16	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	18	13,14,15,16	21	19	O2b*
H434	1	16	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	19	13,14,16	21	19	O2b*
H435	1	16	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	19	16,17	21	20	O2b*
H436	1	16	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	20	16,17	21	20	O2b*
H437	1	16	10,18	12	14	29	23	10	13	13	14	13	12	13	25	18	32	15	18	13,14,16	21	20	O2b*
H438	1	16	10,18	12	14	29	23	10	13	13	14	13	12	14	26	18	30	15	17	13,16	21	20	O2b*
H439	1	16	10,18	12	14	29	23	10	13	13	14	13	13	13	25	18	31	15	17	13,14,16,17	21	20	O2b*
H440	1	16	10,18	12	14	29	23	10	13	13	14	13	13	13	25	18	32	15	16	13,14,16,17	21	20	O2b*
H441	1	16	10,18	12	14	29	23	10	13	13	15	13	12	13	25	18	30	15	16	13,14,16,17	21	20	O2b*
H442	1	16	10,18	12	14	29	23	10	14	13	14	13	12	13	25	18	32	15	18	13,14,16,17	21	20	O2b*
H443	1	16	10,18	12	14	29	23	10	14	13	15	13	12	13	25	18	31	15	19	13,14,16,17	21	20	O2b*
H444	2	16	10,18	12	14	29	24	10	13	13	14	13	12	13	25	18	31	15	16	13,14,16,17	22	20	O2b*
H445	1	16	10,18	12	16	31	22	10	13	13	14	13	11	13	25	18	32	15	18	13,14,16,17	22	20	O2b*
H446	1	16	10,18	13	14	29	23	10	14	13	14	13	11	13	25	18	29	15	19	13,16,17	21	20	O2b*
H447	1	16	10,18	13	14	30	23	10	13	13	14	13	11	13	25	18	30	16	18	14,16,17	20	20	O2b*
H448	1	16	10,19	12	13	28	23	10	13	12	14	13	13	13	25	18	30	14	17	13,14,16	22	20	O2b*
H449	1	16	10,19	12	13	28	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16,18	21	20	O2b*
H450	1	16	10,19	12	13	28	23	10	13	13	14	13	12	14	25	18	29	14	17	13,14,16,17	20	20	O2b*
H451	1	16	10,19	12	13	28	23	10	13	13	14	13	13	13	25	18	31	15	18	13,14,16,17	21	20	O2b*
H452	1	16	10,19	12	13	28	24	10	13	13	14	13	12	14	26	18	33	15	19	12,14,17	21	20	O2b*
H453	1	16	10,19	12	14	29	23	10	11	14	16	10	12	15	27	20	34	15	15	12,14,16	21	20	C3
H454	1	16	10,19	12	14	29	23	10	13	13	14	13	11	13	25	18	32	15	17	13,14,16,17	21	20	O2b*
H455	1	16	10,19	12	14	29	23	10	13	13	14	13	12	12	25	18	32	15	17	13,14,16,17	21	20	O2b*
H456	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	24	18	31	15	17	13,14,16,17	21	20	O2b*
H457	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	30	15	17	13,14,16,17	21	20	O2b*
H458	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	30	15	17	13,16,17	21	20	O2b*
H459	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16	21	20	O2b*
H460	3	16	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16,17	21	20	O2b*
H461	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16,17	22	20	O2b*
H462	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	17	16,17	21	20	O2b*
H463	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	32	15	17	13,14,16,17	21	20	O2b*
H464	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	32	15	18	13,14,16,17	22	20	O2b*
H465	1	16	10,19	12	14	29	23	10	13	13	14	13	12	13	26	18	30	15	17	13,14,16,17	21	20	O2b*
H466	2	16	10,19	12	14	29	23	10	13	13	14	13	12	14	25	18	31	15	18	13,14,16,17	21	20	O2b*
H467	1	16	10,19	12	14	29	23	10	13	13	14	13	12	15	25	18	32	15	18	13,14,17	20	20	O2b*
H468	1	16	10,19	12	14	29	23	10	13	13	14	13	13	13	25	18	31	15	18	13,14,16,17	21	20	O2b*

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H469	1	16	10,19	12	14	29	23	10	13	13	14	13	13	13	25	18	32	15	17	13,14,16	22	20	O2b*
H470	1	16	10,19	12	14	29	23	10	13	13	14	13	13	13	25	18	32	15	18	16	21	20	O2b*
H471	1	16	10,19	12	14	29	23	10	13	13	14	13	13	14	25	18	30	15	17	13,14,15,16	21	20	O2b*
H472	1	16	10,19	12	14	29	23	10	13	13	14	13	14	14	25	18	30	15	18	13,14,16,17	21	20	O2b*
H473	1	16	10,19	12	14	29	23	10	14	13	14	13	12	13	25	18	31	15	17	13,14,16,17	22	20	O2b*
H474	1	16	10,19	12	14	29	23	10	15	13	14	13	13	13	25	18	30	15	18	13,14,16,17	21	20	O2b*
H475	1	16	10,19	12	14	30	22	10	13	13	14	13	12	14	25	18	30	15	18	13,15,16,17	20	21	O2b1
H476	1	16	10,19	12	14	30	23	10	11	15	16	10	12	15	27	20	32	15	16	12,14,16	21	20	C3
H477	1	16	10,19	12	14	30	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,15,17	21	20	O2b*
H478	1	16	10,19	12	14	30	23	10	13	13	14	13	12	13	25	18	32	15	17	13,14,16,17	21	20	O2b*
H479	1	16	10,19	12	14	30	24	10	11	15	14	10	11	16	27	21	31	15	15	12,14,16	21	19	C3
H480	1	16	10,19	12	14	30	24	10	13	13	14	13	12	13	25	18	31	15	17	12,14,16,17	21	20	O2b*
H481	1	16	10,19	13	14	29	23	10	13	13	14	13	12	13	25	18	31	15	18	13,14,16,17	22	20	O2b*
H482	1	16	10,20	12	14	29	23	10	13	13	14	13	11	13	25	18	30	15	18	13,14,16,17	21	20	O2b*
H483	1	16	10,20	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16	22	20	O2b*
H484	1	16	10,20	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16,17	21	20	O2b*
H485	1	16	10,20	12	14	29	23	10	14	12	14	12	13	16	24	19	31	15	19	14,17	20	21	O2b*
H486	1	16	10,20	12	14	29	24	10	11	15	14	10	11	16	27	22	32	15	16	12,14,16	21	20	C3
H487	1	16	10,20	12	14	30	22	10	13	13	14	13	11	14	26	18	31	15	19	13,14,17	20	22	O2b1
H488	1	16	10,20	12	14	30	24	10	11	15	14	10	11	16	27	21	31	15	14	12,14,16	21	19	C3
H489	1	16	10,20	12	14	30	24	10	11	15	14	10	11	16	27	21	31	15	15	12,14,16	21	19	C3
H490	1	16	10,21	12	14	29	23	10	13	13	14	13	12	13	25	18	30	15	16	13,14,16	22	21	O2b*
H491	1	16	10,21	12	14	29	25	10	11	15	14	10	13	17	27	21	32	15	16	12,14,16	21	20	C3
H492	1	16	10,21	12	14	29	25	10	11	15	14	10	13	18	26	21	32	15	17	12,14,16	21	20	C3
H493	1	16	10,21	13	14	29	22	10	13	13	14	13	12	14	25	18	33	15	19	13,15,16	20	21	O2b1
H494	1	16	11,11	12	12	27	23	11	11	13	14	10	13	17	27	22	28.2	15	17	12,14,16	21	20	C3
H495	1	16	11,11	12	13	28	22	10	14	12	14	10	11	11	25	18	31	15	19	13,15,16	21	21	O3a1c
H496	1	16	11,11	12	13	28	22	10	14	12	14	10	12	11	26	18	31	15	17	13,14,15	20	21	O3a1c
H497	1	16	11,11	12	13	28	22	10	14	12	15	10	11	11	26	18	31	15	17	12,13,15	20	21	O3a1c
H498	1	16	11,11	12	13	28	22	10	14	12	15	10	11	12	26	18	30	16	17	13,15	20	21	O3a1c
H499	1	16	11,11	12	13	29	22	10	14	12	14	10	12	11	26	18	31	15	17	13,14,15	20	21	O3a1c
H500	1	16	11,11	13	13	28	22	10	14	12	15	10	12	11	26	18	34	15	19	13,15	21	21	O3a1c
H501	1	16	11,12	12	13	28	22	10	14	12	15	10	11	11	26	19	32	15	18	13,15	20	21	O3a1c
H502	1	16	11,13	12	13	29	22	10	14	13	14	10	11	14	26	19	28	15	20	13	21	21	N
H503	1	16	11,13	12	13	29	22	10	14	13	14	10	11	14	26	19	29	15	18	12,13	21	21	N
H504	1	16	11,14	12	14	30	25	10	11	13	14	11	10	13	24	20	33	16	16	13,15,16	23	22	R
H505	1	16	11,16	12	12	27	23	10	11	14	14	10	12	17	25	21	28	16	15	12,14,17	21	20	C3
H506	1	16	11,16	12	12	27	23	11	11	12	14	10	13	17	26	22	29	16	18	12,13,14,16	21	20	C3
H507	1	16	11,16	12	14	29	22	10	11	15	14	10	12	17	26	21	29	15	16	12,13,17	21	20	C3
H508	1	16	11,16	13	13	29	23	10	11	14	14	10	12	16	26	21	29	14	13	12,16	21	20	C3
H509	1	16	11,16	13	13	29	23	10	11	14	14	10	12	16	26	21	29	14	13	12,15,16	22	20	C3
H510	1	16	11,16	13	13	29	23	10	11	15	14	10	12	16	26	21	29	14	13	12,15,16	21	20	C3
H511	1	16	11,17	12	12	27	23	11	11	13	14	10	13	16	26	22	29	16	16	12,14,16	21	20	C3
H512	1	16	11,17	12	13	28	23	10	13	13	14	13	12	13	25	18	31	16	17	13,14,16	21	20	O2b*
H513	1	16	11,17	12	14	29	23	10	11	13	14	10	13	17	27	21	28	15	18	14,16	21	20	C3d
H514	1	16	11,17	12	14	29	23	10	13	13	14	13	12	13	25	18	30	15	18	13,14,16,17	21	20	O2b*
H515	1	16	11,17	12	14	29	23	10	13	13	14	13	13	14	25	18	30	15	18	13,14,17,18	22	20	O2b*

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H516	1	16	11,17	13	13	29	23	10	11	15	14	10	11	17	26	21	30	15	16	12,14,15	21	20	C3
H517	1	16	11,17	13	13	29	23	10	11	15	14	10	12	16	25	21	30	15	16	13,16,17	21	20	C3
H518	1	16	11,17	13	14	30	23	10	11	15	15	10	11	16	26	21	30	15	16	14	22	20	C3
H519	1	16	11,17	13	14	31	23	10	11	14	14	10	11	16	27	21	30	15	16	12,14,16	21	20	C3
H520	1	16	11,17	14	13	29	23	10	11	14	14	10	11	17	26	22	32	15	16	12,14,16	21	20	C3
H521	1	16	11,18	12	13	28	23	10	13	13	14	13	13	13	25	18	29	15	17	14,16,17	22	20	O2b*
H522	1	16	11,19	12	13	29	22	10	13	13	14	13	11	15	25	18	29	15	19	13,15,16,17	20	20	O2b1
H523	1	16	11,19	13	13	29	22	10	13	13	14	14	11	15	25	18	29	16	16	15,16	20	21	O2b1
H524	1	16	11,19	13	14	30	23	10	11	14	14	10	10	14	25	21	30	15	15	12,13,14,17	21	20	C3
H525	1	16	11,20	12	14	28	23	10	13	13	14	13	12	14	25	18	30	15	17	13,14,15,17	21	20	O2b*
H526	1	16	11,20	13	14	30	23	10	11	14	14	10	12	16	25	20	31	15	16	12,13,14,17	21	20	C3
H527	1	16	12,14	12	12	28	23	10	14	12	14	10	12	12	25	19	33	16	18	13,14	22	22	O1a*
H528	1	16	12,14	12	12	28	23	10	14	12	14	10	13	14	24	19	35	17	17	13,14	21	22	O1a*
H529	1	16	12,15	12	12	28	23	10	14	12	14	10	12	14	24	19	34	16	16	13,14	20	23	O1a*
H530	1	16	12,17	12	12	28	23	10	12	12	16	10	12	14	24	19	34	15	18	13,15	19	21	O3a2c1*
H531	1	16	12,17	12	12	28	23	10	13	12	14	10	11	12	23	18	31	15	16	12,13,14,15	20	21	O3a2c*
H532	1	16	12,17	13	12	28	23	10	11	15	14	10	12	18	26	21	29	15	13	12,14,15	21	19	C3
H533	1	16	12,17	13	12	28	25	10	13	12	14	10	13	13	24	20	31	14	18	13,14,16,18	21	22	O3a1c
H534	1	16	12,17	13	13	29	23	10	11	14	14	10	10	16	27	21	28	15	16	13	22	21	C3
H535	1	16	12,17	13	13	29	23	10	11	14	14	10	11	16	26	20	30	15	13	12,14,16	23	19	C3
H536	1	16	12,17	13	13	29	23	10	12	14	14	10	12	17	28	22	30	15	20	12,13,15,16	21	21	C3
H537	1	16	12,17	14	13	29	23	10	11	14	14	10	12	16	28	22	30	15	18	12,13,14,16	21	20	C3
H538	1	16	12,18	10	12	28	25	10	14	12	15	11	12	12	23	20	29	15	16	12,13,14,17	20	20	O3a2c1a*
H539	1	16	12,18	12	12	27	24	10	13	13	14	10	11	12	24	20	27	14	18	13,14,18	22	20	O3a1c
H540	1	16	12,18	12	12	27	25	10	13	12	14	10	12	12	24	18	33	13	18	13,14,17,18	21	21	O3a1c
H541	1	16	12,18	12	12	28	24	10	13	12	15	10	11	13	24	19	34	15	18	13,14,16,17	22	21	O3a1c
H542	1	16	12,18	12	12	28	25	10	14	12	14	10	12	12	24	20	31	14	18	13,14,15,17	21	21	O3a1c
H543	1	16	12,18	13	12	28	24	10	11	14	14	10	14	16	28	22	30	15	18	12,13,14,16	22	20	C3
H544	1	16	12,18	13	12	29	25	10	13	12	14	10	14	13	24	19	33	14	17	13,14,16,19	21	21	O3a1c
H545	1	16	12,18	13	13	29	22	11	11	14	14	10	13	16	27	22	30	15	18	12,13,14,16	21	20	C3
H546	1	16	12,18	13	13	29	23	10	11	14	14	10	11	15	28	23	30	15	19	12,13,14	21	20	C3
H547	1	16	12,18	13	13	29	23	10	11	14	14	10	12	16	27	22	30	15	18	12,13,14,16	21	20	C3
H548	1	16	12,18	13	13	29	23	10	11	14	14	10	12	16	28	22	30	15	19	12,13,14,16	21	20	C3
H549	1	16	12,18	13	13	29	23	10	11	14	14	10	13	15	27	21	30	15	16	12,13,15,17	22	20	C3
H550	1	16	12,18	13	13	29	23	11	11	14	14	10	11	15	28	22	31	15	18	12,13,14,16	21	20	C3
H551	1	16	12,18	13	13	29	23	11	11	14	14	10	11	16	28	22	31	15	18	12,13,14,16	21	20	C3
H552	1	16	12,18	13	13	29	24	10	11	14	14	10	12	16	28	22	30	16	19	12,13,14,16	21	20	C3
H553	1	16	12,18	13	13	30	23	10	11	14	14	10	12	16	29	22	30	15	18	12,13,14,16	21	20	C3
H554	1	16	12,18	13	14	30	23	10	11	14	14	10	12	16	28	22	29	15	17	12,13,14,16	21	20	C3
H555	1	16	12,19	10	12	28	25	10	14	12	15	11	12	12	23	20	29	15	14	12,13,14,16	20	21	O3a2c1a*
H556	1	16	12,19	10	12	28	25	10	14	12	15	11	13	12	23	20	29	15	15	12,13,14,16	20	21	O3a2c1a*
H557	1	16	12,19	11	13	30	25	10	13	12	13	10	10	12	25	19	30	14	18	13,16,17	21	21	O3a1c
H558	1	16	12,19	12	12	28	25	10	13	12	14	10	13	13	24	19	32	14	17	13,14,18	20	20	O3a1c
H559	1	16	12,19	12	12	29	25	10	13	12	14	10	13	13	24	19	32	14	20	13,15,17	22	21	O3a1c
H560	1	16	12,19	12	12	29	26	10	13	12	14	10	12	12	24	20	32	14	17	13,15,16,18	21	21	O3a1c
H561	1	16	12,19	12	13	29	23	10	11	14	14	10	12	16	28	22	31	15	18	12,13,14,16	21	20	C3
H562	1	16	12,19	12	13	29	25	10	13	12	15	10	11	14	23	19	31	15	19	13,15,16,17	22	20	O3a1c

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H563	1	16	12,19	12	13	30	25	10	13	12	15	10	10	12	24	19	32	14	17	13,15,17	21	22	O3a1c
H564	1	16	12,21	12	12	28	25	10	13	12	14	10	12	12	24	18	33	14	18	14,15,18	21	19	O3a1c
H565	1	16	13,13	12	12	28	23	10	14	13	14	10	12	13	25	18	30.2	12	17	13,14	20	21	O1a1
H566	1	16	13,13	12	12	29	23	10	14	13	14	10	12	14	25	18	30	17	15	12,13,14	19	21	O1a1
H567	1	16	13,13	12	12	29	25	10	13	12	14	10	13	12	24	20	32	14	19	13,14,15,18	21	22	O3a1c
H568	1	16	13,13	12	12	30	24	11	14	13	14	10	11	14	25	18	31	15	14	12,13,14	19	20	O1a1
H569	1	16	13,13	12	13	29	23	11	14	13	14	10	12	13	25	18	29	16	15	11,13,14	21	21	O1a1
H570	1	16	13,17	12	12	29	25	10	13	12	15	10	13	11	23	20	35	15	18	13,14,15	20	22	O3a2b
H571	1	16	13,17	12	12	29	25	10	13	12	15	10	13	11	23	20	35	15	18	13,14,15	20	23	O3a2b
H572	1	16	13,17	12	12	30	24	10	14	12	14	10	12	12	24	19	31	14	17	12,13,14,15	23	21	O3a1c
H573	1	16	13,18	12	12	28	24	11	13	12	14	10	13	13	26	19	32	14	17	13,14,16	22	20	O3a1c
H574	1	16	13,18	12	13	29	23	10	12	13	14	10	12	14	21	19	29	15	17	14,15	22	21	O2*
H575	1	16	13,18	13	13	29	23	10	11	14	14	10	12	16	27	22	30	15	18	12,14,16	21	20	C3
H576	1	16	13,18	13	14	30	22	10	11	14	14	10	12	16	28	22	30	15	18	12,13,15,16	21	20	C3
H577	1	16	13,19	12	12	28	24	11	13	12	14	10	12	13	26	19	31	14	17	13,14,15,16	21	20	O3a1c
H578	1	16	13,19	12	12	29	25	11	13	12	14	11	13	12	24	19	31	14	19	11,13,16,18	20	20	O3a1c
H579	1	16	13,20	12	12	28	23	11	13	12	14	10	12	13	26	19	31	14	18	13,14,15,16	22	20	O3a1c
H580	1	16	13,20	12	12	28	23	11	13	12	14	10	12	14	26	19	31	14	17	13,14,15,16	22	20	O3a1c
H581	1	16	13,20	12	12	28	23	11	13	12	14	10	13	13	26	19	30.1	14	18	13,14,15,16	22	20	O3a1c
H582	1	16	13,20	12	14	31	23	7	13	12	14	10	12	13	23	19	30	15	17	12,16	23	19	O3a1
H583	1	16	14,17	12	12	27	23	10	13	12	14	10	11	13	24	20	31	14	19	13,14	22	20	O3a1c
H584	1	16	14,17	12	12	28	24	10	13	12	14	10	11	13	24	20	26	14	19	12,13,14,17	21	21	O3a1c
H585	1	16	14,17	12	13	29	23	10	11	14	14	10	13	15	26	18	31	15	14	17	22	21	D
H586	1	16	14,17	12	13	29	26	10	13	12	14	10	11	12	24	20	27	15	18	13,14,17	24	21	O3a1c
H587	1	16	14,18	12	12	27	24	10	13	12	14	10	11	13	25	20	29	14	17	13,14	21	21	O3a1c
H588	1	16	14,18	12	12	27	25	11	13	12	14	10	11	12	24	19	27	13	18	13,15,17	23	21	O3a1c
H589	1	16	14,18	12	12	28	25	10	13	12	14	10	11	12	24	20	27	14	19	13,15,18	24	21	O3a1c
H590	1	16	14,18	12	12	29	26	10	13	12	14	10	11	12	24	19	27	14	19	13,18	23	21	O3a1c
H591	1	16	14,18	12	13	29	23	10	11	14	14	9	12	15	26	18	31	15	17	17	21	21	D
H592	1	16	14,19	12	14	31	24	10	12	12	15	10	11	12	22	21	33	13	17	11,15,16,17	21	21	O3a1
H593	1	16	14,19	13	12	28	24	10	13	13	15	10	12	11	23	20	31	15	16	12,13,15,16	21	22	O3a2*
H594	1	16	14,20	13	12	30	23	10	13	12	14	11	11	11	24	20	31	15	19	12,14,15,17	21	22	O3a2*
H595	1	16	14,20	13	12	30	23	10	13	12	15	10	11	11	23	20	33	16	16	12,14,15,16	21	21	O3a2*
H596	1	16	14,20	15	13	29	23	10	11	15	14	10	11	17	25	21	28	15	17	13,14,16	21	20	C3
H597	1	16	14,21	13	12	28	26	10	13	12	14	9	11	11	24	20	31	15	16	12,14,15,16	21	19	O3a2*
H598	1	16	14,21	13	12	30	24	10	13	12	14	10	11	11	24	20	31	16	16	12,14,15,16	23	21	O3a2*
H599	1	16	14,21	13	12	30	25	10	14	13	14	11	11	11	24	20	31	15	17	12.3,14,15,17	21	22	O3a2*
H600	1	16	15,15	12	14	31	24	10	12	12	15	10	12	12	22	21	32	13	17	12,15,16	21	22	O3a1
H601	1	16	15,18	12	13	30	25	11	12	12	15	10	12	12	22	21	34	14	16	12,15,16,17	22	21	O3a1
H602	1	16	15,19	13	12	30	24	10	13	12	15	10	11	11	24	20	32	15	16	12,14,15	22	22	O3a2*
H603	1	16	15,19	13	12	30	24	10	13	12	15	10	11	11	24	20	33	15	16	15,16	21	21	O3a2*
H604	1	16	15,20	13	12	28	24	10	13	12	15	10	11	11	23	20	29	15	16	12,14,15,16	22	20	O3a2*
H605	1	16	15,20	13	12	29	24	10	13	12	15	10	11	11	24	20	33	15	16	12,14,15,16	22	20	O3a2*
H606	1	16	15,20	13	12	29	24	11	13	12	15	10	11	11	23	20	31	16	16	12,14,15,16	21	21	O3a2*
H607	1	16	15,21	13	12	29	24	10	13	12	15	10	11	11	23	20	31	15	16	12,14,15,17	21	20	O3a2*
H608	1	16	15,21	13	12	29	24	10	13	12	15	10	11	11	24	20	31	15	16	12,14,15,16	21	21	O3a2*
H609	1	16	15,21	13	12	29	24	10	13	12	15	10	11	11	24	20	31	15	18	12,14,15	21	21	O3a2*

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H610	1	16	15,22	13	12	29	24	10	13	12	15	10	11	11	24	20	31	15	16	12,14,15,16	21	21	O3a2*
H611	1	16	15,22	13	12	30	24	10	13	12	15	10	11	11	23	20	31	16	15	12,14,15,16	21	21	O3a2*
H612	1	16	15,22	13	12	30	24	10	13	12	15	10	11	11	24	20	31	15	18	12,14,15,16	23	22	O3a2*
H613	1	16	15,22	13	12	30	24	10	13	12	15	10	11	13	24	20	31	15	16	12,14,16	21	21	O3a2*
H614	1	16	16,21	13	12	29	24	10	13	12	15	10	12	11	23	20	31	15	16	12,14,15,17	21	20	O3a2*
H615	1	16	19,20	12	14	29	23	10	11	14	14	10	12	15	28	21	28	15	18	12,14,16	20	19	C3d
H616	1	17	9,16	12	14	30	25	11	11	13	14	10	14	14	24	17	34	17	17	15,16,17	24	20	D
H617	1	17	10,17	12	13	28	23	11	13	14	14	13	12	13	26	18	30	16	17	14,15,17	21	20	O2b*
H618	1	17	10,17	12	14	29	23	11	13	13	14	13	11	13	25	18	33	15	19	13,14,17	21	20	O2b*
H619	1	17	10,18	12	13	28	23	10	13	13	14	13	12	13	25	18	30	15	17	13,14,15,17	22	20	O2b*
H620	1	17	10,18	12	13	28	23	10	13	13	14	13	12	14	25	18	33	15	17	13,15,16,17	23	20	O2b*
H621	1	17	10,18	12	13	28	23	10	14	13	14	13	12	13	26	18	31	16	17	13,14,15,17	21	20	O2b*
H622	1	17	10,18	12	14	29	23	10	13	13	14	13	12	14	25	18	30	15	18	13,16,17	21	20	O2b*
H623	1	17	10,19	12	14	28	23	10	11	16	15	10	12	16	28	20	33	15	16	12,14,17	21	20	C3
H624	1	17	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	17	13,14,16	21	20	O2b*
H625	1	17	10,19	12	14	29	23	10	13	13	14	13	12	13	25	18	31	15	18	13,14,17	21	21	O2b*
H626	1	17	10,19	12	14	29	23	10	13	13	14	13	12	14	25	18	31	15	18	13,14,16,17	20	20	O2b*
H627	1	17	10,19	12	14	30	23	10	13	13	14	13	11	14	25	18	30	16	19	13,14,15,17	21	20	O2b*
H628	1	17	10,19	12	14	30	23	10	13	13	14	13	12	13	26	18	30	15	17	13,14,16,17	21	20	O2b*
H629	1	17	10,20	12	14	29	24	10	11	15	14	10	11	17	26	21	33	15	17	12,14,16	22	20	C3
H630	1	17	11,16	12	12	28	24	11	11	13	14	10	12	17	26	23	29	15	17	12,14,16	22	20	C3
H631	1	17	11,19	12	13	30	25	10	13	12	15	10	12	13	22	19	33	14	18	13,14,15	22	22	O3a1c
H632	1	17	11,20	12	13	29	25	10	13	13	15	10	12	13	23	19	33	14	19	14	21	21	O3a1c
H633	1	17	11,20	12	13	30	25	10	13	12	15	10	12	13	23	19	34	14	18	13,14	21	21	O3a1c
H634	1	17	11,20	12	14	31	25	9	13	12	14	10	12	14	23	19	34	14	17	13,14	22	21	O3a1c
H635	1	17	11,21	12	13	30	24	10	13	13	14	10	11	12	23	19	34	14	17	13,14	22	21	O3a1c
H636	1	17	12,13	12	13	28	23	10	12	12	15	10	14	13	25	18	32	15	20	13,14,16	19	21	O3a2c1*
H637	1	17	12,16	13	14	31	27	10	11	13	14	10	14	18	24	17	31	15	17	15,17	21	21	D
H638	1	17	12,17	12	11	27	24	10	13	12	14	10	12	13	24	20	31	15	18	13,14,18	21	21	O3a1c
H639	1	17	12,17	12	12	28	23	10	13	12	14	10	12	12	24	19	32	14	19	13,14,15	21	21	O3a1c
H640	1	17	12,17	12	12	28	24	10	13	12	14	10	12	12	24	19	32	14	18	13,14,15	21	21	O3a1c
H641	1	17	12,17	12	12	29	24	10	13	12	14	10	12	12	24	19	31	14	18	13,14,15	21	22	O3a1c
H642	1	17	12,17	13	14	30	23	10	11	14	14	10	13	16	29	22	32	15	18	12,13,14,16	21	20	C3
H643	1	17	12,18	12	12	28	24	11	13	12	14	10	11	11	24	20	32	14	18	13,16,18	22	20	O3a1c
H644	1	17	12,18	14	13	29	24	10	13	12	14	10	12	12	23	19	29	14	19	13,16,17	21	20	O3a1c
H645	1	17	12,19	10	12	28	25	10	14	12	15	11	12	12	23	20	28	15	15	12,13,14,16	20	21	O3a2c1a*
H646	1	17	12,19	10	12	28	25	10	14	12	15	11	12	12	23	20	29	15	15	12,13,14,16	20	21	O3a2c1a*
H647	1	17	12,19	10	12	28	25	10	14	12	15	11	13	12	23	20	30	15	15	12,14,17	20	21	O3a2c1a*
H648	1	17	12,19	12	12	28	25	10	13	12	14	10	12	13	24	20	30	14	19	12,14,15,17	21	21	O3a1c
H649	1	17	12,20	10	12	28	25	10	14	12	15	11	12	12	23	20	29	15	15	11,13,14,16	20	21	O3a2c1a*
H650	1	17	12,20	12	12	28	24	10	13	12	14	10	12	12	24	19	30	15	17	12,14,15,16	24	20	O3a1c
H651	1	17	12,20	12	12	28	25	10	13	12	14	10	11	12	24	18	28	15	19	13,15,16	25	21	O3a1c
H652	1	17	12,20	12	12	29	24	10	13	12	14	10	11	12	24	20	29	14	18	13,14,17	21	21	O3a1c
H653	1	17	13,15	12	14	31	24	10	11	13	14	11	12	17	25	17	27	15	16	17	21	20	D
H654	1	17	13,15	13	13	29	23	10	11	14	14	10	11	16	26	21	31	15	16	13,15	20	20	C3
H655	1	17	13,16	12	13	29	25	10	11	13	14	10	12	18	25	17	27	16	16	17,18	21	20	D
H656	1	17	13,16	12	13	30	25	10	11	13	14	10	12	17	24	17	27	15	16	17	21	20	D

Table 14. Cont.

Haplotype	N	DYS-19	DYS-385	DYS-388	DYS-389I	DYS-389II	DYS-390	DYS-391	DYS-392	DYS-393	DYS-437	DYS-438	DYS-439	DYS-446	DYS-447	DYS-448	DYS-449	DYS-456	DYS-458	DYS-464	DYS-635	GATA H4.1	Haplogroup
H657	1	17	13,16	12	13	30	25	10	11	13	14	10	13	18	25	17	28	16	16	17,18	21	20	D
H658	1	17	13,16	12	13	30	26	10	11	13	14	10	12	19	25	17	27	16	16	17	21	19	D
H659	1	17	13,16	12	13	31	24	10	11	13	14	10	12	18	25	17	28	16	15	17	21	20	D
H660	1	17	13,16	12	14	31	24	10	11	13	14	10	11	17	25	17	27	15	16	17	21	20	D
H661	1	17	13,17	12	12	30	24	10	14	12	14	10	11	12	24	19	32	14	18	13,14,15	22	21	O3a1c
H662	2	17	13,18	12	13	30	25	11	11	13	14	10	12	19	25	19	33	15	15	16,18	21	20	D
H663	1	17	13,18	12	14	31	25	10	11	13	14	10	12	18	25	17	33	15	16	15,17	21	21	D
H664	1	17	13,19	12	13	29	25	10	13	14	14	10	11	12	25	20	33	13	17	12,13,16	22	21	O3a1c
H665	1	17	13,19	12	14	31	25	10	13	12	15	10	13	12	25	19	32	15	18	15,17	23	20	O3a1c
H666	1	17	13,20	12	13	30	26	10	11	13	14	10	12	15	26	17	29	15	16	17	22	21	D
H667	1	17	13,21	12	12	29	25	11	13	12	14	10	11	12	24	19	31	14	18	13,15,17	21	21	O3a1c
H668	1	17	14,17	12	12	27	24	10	13	12	14	10	12	12	24	19	27	14	18	13,15,18	23	21	O3a1c
H669	1	17	14,18	12	12	27	24	10	13	12	14	10	12	12	24	19	27	14	18	13,15,18	23	22	O3a1c
H670	1	17	14,18	12	12	27	25	10	13	12	14	10	13	12	24	19	27	14	18	13,15,18	23	21	O3a1c
H671	1	17	14,18	12	12	27	25	11	13	12	14	10	11	12	24	19	27	14	21	12,13,15,18	22	21	O3a1c
H672	1	17	14,18	12	12	27	25	11	13	13	14	10	12	12	24	19	27	14	17	13,15,18	23	21	O3a1c
H673	1	17	14,18	12	12	28	25	11	13	12	14	10	12	12	24	19	27	14	18	13,15,19	24	21	O3a1c
H674	1	17	14,19	13	12	27	25	10	13	12	14	10	12	12	24	19	27	14	17	13,15,18	22	21	O3a1c
H675	1	17	14,19	13	12	29	24	10	13	12	14	10	12	11	24	20	32	15	16	12,13,16	22	21	O3a2*
H676	1	17	14,20	13	12	29	24	10	14	12	15	10	11	11	24	20	31	15	16	12,14,16	21	21	O3a2*
H677	1	17	14,20	13	12	30	24	10	13	12	14	10	11	12	24	20	32	15	18	12,14,15,16	21	21	O3a2*
H678	1	17	14,20	13	12	30	24	11	13	12	14	10	11	11	24	20	31	15	16	13,14,15,16	21	21	O3a2*
H679	1	17	14,20	13	13	29	24	10	13	13	15	10	13	11	23	20	32	15	15	12,14,16	21	20	O3a2*
H680	1	17	14,21	13	12	29	24	10	13	12	14	10	11	11	24	0	31	15	16	14,16	20	20	O3a2*
H681	1	17	14,21	13	12	29	24	10	13	13	14	10	11	11	24	0	31	15	16	14,16	20	20	O3a2*
H682	1	17	15,19	12	14	30	24	10	12	12	15	10	12	12	22	20	32	14	17	12,15,16,17	21	21	O3a1
H683	1	17	15,20	13	12	29	24	10	13	12	14	10	12	11	24	20	31	15	16	12,14,15,16	21	21	O3a2*
H684	1	17	15,20	13	12	29	24	10	13	12	15	10	11	11	24	19	31	15	17	12,14,15,16	21	21	O3a2*
H685	1	17	15,20	13	12	30	24	10	13	12	15	10	11	13	24	20	31	15	16	11,14,16	21	21	O3a2*
H686	1	17	15,20	13	12	30	24	10	13	12	15	10	12	11	24	20	31	15	16	13,14,15,16	21	21	O3a2*
H687	1	17	15,21	13	12	29	24	10	13	12	15	10	11	11	24	20	32	15	18	12,13,15,16	22	21	O3a2*
H688	1	17	15,21	13	12	30	24	10	13	12	14	10	11	11	24	20	30	15	17	12,14,15,16	21	21	O3a2*
H689	1	17	16,20	13	12	30	24	10	13	12	15	10	12	11	24	20	29	15	16	12,14,15,16	21	21	O3a2*
H690	1	17	16,21	13	12	29	24	10	13	12	15	10	11	11	24	20	31	15	17	12,14,15,16	21	21	O3a2*
H691	1	17	17,20	13	13	30	24	10	13	12	15	10	13	11	24	20	32	15	17	12,14,15,16	21	21	O3a2*
H692	1	18	14,14	13	12	29	24	10	13	12	15	10	11	11	24	20	31	15	19	12,15,16	22	22	O3a2*
H693	1	16,17	10,17	12	13	28	23	10	13	13	14	13	12	12	25	18	30	16	18	16	21	20	O2b*

Table 15. Number of haplotypes and diversity value for 22 Y-STRs in 706 unrelated Korean males

	9 Y-STRs haplotype[a]	11 Y-STRs haplotype[b]	17 Y-STRs haplotype[c]	22 Y-STRs haplotype[d]
No. of haplotypes	485	558	657	693
No. of unique haplotypes	396	485	627	682
Discriminatory capacity (%)	68.7	79.0	93.1	98.2
Haplotype diversity (±SD)	0.9966 ± 0.0005	0.9982 ± 0.0004	0.9995 ± 0.0002	0.9999 ± 0.0001

[a] The 9 Y-STRs are a combination of DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393.

[b] The 11 Y-STRs are a combination of DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438 and DYS439.

[c] The 17 Y-STRs are a combination of DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and GATA H4.1.

[d] The 22 Y-STRs are DYS19, DYS385a/b, DYS388, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS446, DYS447, DYS448, DYS449, DYS464, DYS456, DYS458, DYS635 and GATA H4.1.

To assess the influence of each Y-STR in haplotype context, the number of haplotypes and haplotype diversities were calculated by adding extra loci to the extended SWGDAM haplotype (Table 16). Among the added loci, DYS458 showed the largest increase in the number of haplotypes and in the haplotype diversity value, although the locus did not exhibit the highest gene diversity (0.7779) among the 10 added loci. DYS449 which showed the highest gene diversities (0.8523), contributed to the increase in haplotype diversity as well. Another highly variant STR, DYS446, increased the haplotype diversity, but DYS448 did not contribute significantly to a haplotype diversity increment.

Table 16. Number of haplotypes and diversity value for each Y-STR haplotype in 706 unrelated Korean males

Haplotype	No. of haplotype	Discriminatory capacity (%)	Haplotype diversity
11 Y-STRs[a]	558	79.0	0.9982
11 Y-STRs+DYS449	612	86.7	0.9992
11 Y-STRs+DYS446	598	84.7	0.9989
11 Y-STRs+DYS458	616	87.3	0.9992
11 Y-STRs+DYS447	583	82.6	0.9986
11 Y-STRs+DYS448	565	80.0	0.9983
11 Y-STRs+DYS635	591	83.7	0.9988
11 Y-STRs+GATA H4.1	579	82.0	0.9985
11 Y-STRs+DYS456	574	81.3	0.9984
11 Y-STRs+DYS388	566	80.2	0.9983
11 Y-STRs+DYS437	568	80.5	0.9983
11 Y-STRs+DYS458+DYS635	639	90.5	0.9994
11 Y-STRs+DYS458+DYS447	630	89.2	0.9994
11 Y-STRs+DYS458+GATA H4.1	631	89.4	0.9993
11 Y-STRs+DYS458+DYS635+DYS447	650	92.1	0.9996
11 Y-STRs+DYS458+DYS635+GATAH4.1	649	91.9	0.9995
11 Y-STRs+DYS458+DYS635+DYS447+GATA H4.1	658	93.2	0.9996
17 Y-STRs[b]	657	93.1	0.9995
17 Y-STRs+DYS449	673	95.3	0.9998
17 Y-STRs+DYS446	670	94.9	0.9996
17 Y-STRs+DYS447	665	94.2	0.9996
17 Y-STRs+DYS388	660	93.5	0.9995
17 Y-STRs+DYS447+DYS388	668	94.6	0.9996
22 Y-STRs[c]	693	98.2	0.9999

[a] The 11 Y-STRs are a combination of DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438 and DYS439.

[b] The 17 Y-STRs are a combination of DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and GATA H4.1.

[c] The 22 Y-STRs are DYS19, DYS385a/b, DYS388, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS446, DYS447, DYS448, DYS449, DYS464, DYS456, DYS458, DYS635 and GATA H4.1.

The DYS458 locus which generated highest increase in haplotype diversity

was added the SWGDAM haplotype combination, and then the effect for the addition of the other markers was also assessed. The addition of DYS458 and DYS635 to the extended SWGDAM haplotype generated the largest increase in the number of haplotypes. Finally, The addition of DYS447, DYS458, DYS635, and GATA H4.1 resulted in a comparable number of haplotypes and haplotype diversities to the AmpFlSTR® Yfiler™ haplotype, despite the fewer loci than with AmpFlSTR® Yfiler™ (15 loci vs. 17 loci) were used.

Meanwhile, adding extra loci to the AmpFlSTR® Yfiler™ loci haplotype was also tested, because the 17-Yfiler loci are widely used in forensic field. Likewise, DYS449, DYS446, DYS447 and DYS388, in order of those, contributed to the increase of haplotype number. The addition of DYS447 and DYS388 showed somewhat increase in the number of haplotype and haplotype diversity but the increase was still lower than that by addition of DYS449 or DYS446.

The Korean 22 Y-STRs haplotypes were classified into Y-chromosomal haplogroups. I did not observe any individuals sharing the same 22 Y-STRs haplotype among different haplogroups, even when 17 Yfiler-loci haplotype was applied. However, when considering minimal haplotypes, three cases of haplotype sharing were observed between haplogroup O2b-SRY_{465} and O2b1-47z and one between haplogroup O3a1c-JST002611 and O3a2c1a-M117. The apportionment of 22 Y-STRs haplotype variability among and within the 21 observed haplogroups was determined by AMOVA; only 38.95%

(p<0.0001) of the genetic variation was attributable to the difference among haplogroups in the Korean population (Table 17).

In separate AMOVA analysis, DYS392, DYS393, DYS437, DYS438, DYS448 and DYS388 loci showed that its variability is highly structured between haplogroups (Table 10). In this context, the combined haplotypes consisting only of DYS392, DYS393, DYS437, DYS438, DYS448 and DYS388 loci were constructed and assessed by AMOVA analysis (Table 17). The analysis revealed 71.15% of the total variation among haplogroups for the combined haplotypes, with 38.95%, 41.52%, and 42.09% of total variation among haplogroups for all 22 Y-STRs, 17 Y-STRs and minimal haplotypes, respectively.

Table 17. Diversity and AMOVA analysis for each Y-STR haplotype

Haplotype	Haplotype diversity	% Variance	
		Among haplogroups	Within haplogroups
DYS392-DYS393-DYS438-DYS437-DYS448-DYS388	0.9317	71.15	28.85
9 Y-STRs[a]	0.9966	42.09	57.91
11 Y-STRs[b]	0.9982	43.71	56.29
17 Y-STRs[c]	0.9995	41.52	58.48
22 Y-STRs[d]	0.9999	38.95	61.05

[a] The 9 Y-STRs are a combination of DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393.

[b] The 11 Y-STRs are a combination of DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438 and DYS439.

[c] The 17 Y-STRs are a combination of DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and GATA H4.1.

[d] The 22 Y-STRs are DYS19, DYS385a/b, DYS388, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS446, DYS447, DYS448, DYS449, DYS464, DYS456, DYS458, DYS635 and GATA H4.1.

To identify the haplogroup affiliation of the combined haplotype (DYS392-DYS393-DYS438-DYS437-DYS448-DYS388) with empirical data, the haplotypes observed frequently were determined for each Korean haplogroup, and I searched for the shared haplotypes, excluding DYS388, in the YHRD release 39 (Feb, 2012)[16] (Table 18). Using the six Y-STRs, most presented haplotypes were classified based on each haplogroup background, except for the non-differentiated haplotype (13-13-13-14-18-12) between haplogroup O2b*-SRY_{465} and O2b1-47z and a haplotype 14-13-10-14-18-12 between haplogroup O1a1-P203 and haplogroup N-M231. However, the addition of DYS390 and DYS389I to the six Y-STR haplotypes could differentiate common haplotypes between the O2b*-SRY_{465} (allele 23-24) and O2b1-47z (most allele 22) and between the N-M231 (allele 13-14) and O1a1-P203 (allele 12), respectively. In the YHRD database search, the matched haplotypes (except DYS388) tended to belong to the same haplogroups as those in Koreans, although some sublineages of haplogroup O3 were not further divided in the YHRD, in comparison to this study. These findings indicate that the six Y-STR loci with lower mutation rates seem to be more strongly structured in the haplogroup background.

Table 18. Representative haplotypes for each haplogroup and the number of matched haplotypes in the YHRD database

Haplotype[a] (392-393-438-437-448-388)	N[b]	%[b]	Haplogroup	YHRD database release 39																			
				DE 1343[c]	D 533[c]	C 498[c]	C3 147[c]	F 7109[c]	Q 226[c]	N 273[c]	O 2321[c]	O1a 96[c]	O1a1 19[c]	O2 629[c]	O2b 555[c]	O2b1 181[c]	O3 1129[c]	O3a 963[c]	O3a2 755[c]	O3a2b 4[c]	O3a2c1 396[c]	O3a2c1a 188[c]	Total 8963[c]
11-13-10-14-17-12	10	55.6	D-M174	88	88	1		1															90
11-13-10-14-19-12	2	11.1		222	208	3		2															227
11-13-11-14-17-12	2	11.1		29	28																		29
11-15-10-14-21-12 (or 13)	23	26.4	C3-M217			115	38																115
11-14-10-14-21-12 (or 13)	19	21.8		2		74	29																76
11-14-10-14-22-13	12	13.8				45	15	1															46
11-13-11-14-0-13	4	4.6				11	5	3															14
14-14-12-14-19-12	7	53.8	Q-M207					20	18														20
15-14-12-14-19-12	3	23.1						8	5														8
14-13-10-14-20-13	5	18.5	N-M231					17		14	2	1					1	1					17
14-13-10-14-19-12 (or 13)	4	14.8						39		27	10	2		5	5								39
15-13-10-14-19-12 (or 13)	4	14.8						15		14													15
16-14-11-14-19-12	3	11.1						10		10													10
14-13-10-14-18-12	3	11.1						60		8	52	44		1									60
14-13-10-14-18-12[f]	13	68.4	O1a1-P203					60		8	52	44	13	1									60
16-13-10-14-18-12[f]	2	10.5						4		1	3	3	2										4
13-13-10-14-18-12	3	30.0	O2*-P31					15			15	1		11									15
13-14-10-14-18-12	2	20.0						45			45			37									45
13-13-13-14-18-12 (or 13)	124	80.0	O2b*-SRY$_{465}$					717			714			398	398	128							717
13-13-13-14-18-12	46	64.8	O2b1-47z					717			714			398	398	128							717
13-12-10-15-21-*12*	6	54.5	O3a1-KL2[h]					12			12						12	10	2				12
12-12-10-15-21-*12*	3	27.3						7			7						7	6					7
13-12-10-14-19-12	24	33.3	O3a1c-JST002611					129			129						129	96	10				129
13-12-10-14-20-12	16	22.2						70			70						70	55	16				70
13-12-10-14-20-*13*	7	16.7	O3a2*-P201					70			70						70	55	16				70
13-12-10-15-20-*13*	25	59.5						121			121						121	93	77	4	1		121
13-12-10-15-20-*12*	3	100	O3a2b-M7					121			121						121	93	77	4	1		121
13-12-9-15-19-12	3	33.3	O3a2c*-P164[h]					4			4						4	4	4				4
13-13-10-15-19-12	2	22.2						27		1	22						22	18	15				27
12-12-10-15-19-12	36	52.9	O3a2c1-M134					122			122						122	122	115		77		122
12-13-10-15-19-12	9	13.2						32			30						30	30	28		20		32
14-12-11-15-20-10	53	58.9	O3a1c1a-M117					211			211						211	211	202		150	110	211

[a]This is a set of DYS392, DYS393, DYS438, DYS437, DYS448 and DYS388 loci. The searched haplotype in the YHRD database is indicated in bold and the same haplotypes belonging to different haplogroups are underlined. The numbers in parentheses indicated the alternative allele of the DYS388 marker in each haplotype. The DYS388 allele in italics indicates an allele only observed in the corresponding Korean haplogroups.

[b]N and % represent the number of matched haplotypes and the proportion of haplotypes in the corresponding Korean haplogroups, respectively.

[c]The number indicates the number of haplotype entries in the corresponding haplogroups from the searched YHRD database.

IV. Discussion

Y-STR allele and haplotype frequencies, diversity values and their haplogroup memberships were analyzed in 706 Korean males. Atypical alleles, such as unusually short allele, partial insertion/deletion mutations or intermediate-sized variants, appeared in the Korean population and their informative frequencies can provide useful information in forensic practice. These atypical alleles arise from mutation in the flanking sequences which alter the length of the PCR product. Therefore, the designation of an atypical allele is sometimes difficult according to the ISFG recommendations for nomenclature[41] without sequence data, for example ostensible alleles 25, 30.1 and 24 at DYS449. For DYS385, allele designation discrepancies were observed in some allele when different primer pairs were used due to the 4-bp and 8-bp deletion mutations in the flanking region. Considering the ISFG recommendations[41], a primer set of multiplex STR-III may be preferable to a primer set of multiplex STR-I/II in that it reflects only the variations in DYS385 core repeats. Moreover, a primer set of multiplex STR-III may be very useful in developing size-reduced STR products. However, many databases have already been constructed using commercial multiplex kits like Powerplex® Y and AmpFlSTR® Yfiler™. These kits produce identical genotyping results with a primer set of multiplex STR-I/II in the present study. In addition, considering the data of Füredi et al.[54], the PCR product amplified by a primer set of

multiplex STR-III includes a rare thymidine deletion downstream from the core repeat unit. None of the 706 sampled Koreans showed this deletion. Therefore, the selection of primer pairs for criminal investigations or for storing DNA profiles in national DNA databases should be carried out with much more consideration of the purpose of analysis and the possible presence of mutations in flanking sequences.

Deletions of DYS448 and DYS464 were observed in association with the *AZFc* rearrangement. In some reports,[55, 56] it was suggested that the *AZFc* marker DYS464 should not be used for commercial typing because of the possibility of interpretive problems related to inadvertent diagnosis of male infertility and because of its characteristics as a multi-locus STR. The DYS448 null allele has been observed in various populations; 10 out of 1079 in Japanese males[57], 3 of 769 in Nepalese men[58], 3 of 980 in three ethnic groups (Malays, Chinese and Indians) living in Malaysia[59], 7 of 99 in Kalmyk[60], 1 of 326 in Mexican[61], 1 of 247 in Spanish[62], and 2 of each 330 Asians, 985 African Americans, and 1276 Caucasians in these three racial groups studied in the Yfiler™ haplotype database by Applied Biosystems (http://www6.appliedbiosystems. com/yfilerdatabase/). The relatively high frequencies of the DYS448 null allele in Asians suggest giving careful consideration to the use of DYS448 for commercial genotyping and further database construction in Asians.

Currently available Y-STR haplotypes including several commercial kits of

up to 17 markers such as AmpFlSTR® Yfiler™ did not have enough resolution power (~93.1%) for differentiating Korean paternal lineage. Haplotype diversity can be increased and male lineage differentiation can be improved by adding additional simple single-copy Y-STR markers. This study showed that the additions of each DYS458, DYS635 (only to SWGDAM haplotype), DYS449, and DYS446 to the haplotypes increased the discriminatory power and the combined haplotypes of 15 loci, DYS447, DYS458, DYS635, GATA H4.1, and the SWGDAM Y-STRs, was comparable to the haplotypes of 17 loci in the AmpFlSTR® Yfiler™ kit. This may be caused by the higher mutation rates of those Y-STRs relative the others, generating greater allele ranges for each marker, which in turn gives the potential for a large number of Y-STR haplotypes. In addition, separate AMOVA analysis for the each marker showed that the proportion of variation among haplogroups was much lower than that within haplogroups, thus presenting weaker relationships with genetic populations association with the geographic clustering of haplogroups. This result corresponds with a study showing that rapidly mutating Y-STRs would detect considerably less genetic population substructure.[20] Among the reported rapidly mutating Y-STRs, DYS449 was also found to have a weaker relationship with haplogroup affiliation in the Korean population. Therefore, fast-mutating STRs such as DYS449 and DYS458 can increase the discriminatory capacity when added to a certain haplotype context[20, 25] such that the potential for distinguishing between paternal lineages would increase,

thereby enabling high-resolution differentiation[20] in a global population.

In contrast, it was found that genetic and haplotype variations of DYS392, DYS393, DYS437, DYS438, DYS448 and DYS388 with a relatively lower mutation rate were highly structured in a haplogroup background. The fraction of the STR variability was higher among haplogroups than within haplogroups. This may be explained by the fact that STRs with a low mutation rate preserved any signal of the haplogroup history or founder effects and generated a small allele range due to their low mutation rates, thereby maintaining a non-homogenous allele distribution across haplogroups. This finding is consistent with the finding that the distributions of allele frequencies at DYS392 and DYS438 reflect evolutionary lineages.[63] Collectively, Y-STRs with relatively low mutation rates can be more informative in terms of inferring a population or a haplogroup substructure which reflects geographic region, in contrast with fast mutating Y-STRs differentiating paternal lineages.

On the other hand, these atypical alleles were found to have certain haplogroup affiliations even though the STRs have relatively faster mutation rates. The unusually short alleles (17 and 18) at DYS447 and a variant allele 30.2 at DYS449 arose on haplogroups O3a1-KL2 (xO3a1c) and O1a1-P203, and likely define sublineages within those haplogroups, respectively. It should be noted that most intermediate alleles at DYS449 occur on multiple haplogroup backgrounds, whereas the allele 30.2 observed in the Korean samples and in one Japanese sample showed a strong association with a

subclade within haplogroup O1a1-P203, similar to alleles 32.2 and 32.3, which were related with haplogroup A-P97 in Cameroon.[28] Unlike atypical alleles of DYS447 and DYS449, the allele 14.3 at DYS464 might occur within as many as two sublineages in haplogroup N-M231. Haplogroup N is a major clade accompanying its derived SNPs, so a higher resolution analysis using the derived SNP is needed to define the specific nature of the substructure for intermediate allele 14.3 at DYS464. Interestingly, a single haplotype sample with allele 17.2 at DYS458 along with allele 15 at DYS388 was found in this Korean population; the pattern of intermediate allele at DYS458 and alleles with ≥15 repeats at DYS388 is confined to a subclade in haplogroup J1-M267.[27, 64] Additionally, the repeat sequence structure $(GAAA)_{15}AA(GAAA)_2$ was coincident with the intermediate allele at DYS458 in the haplogroup background J1-M267.[27] Overall, while the DYS447, DYS449, DYS458 and DYS464 markers have a higher mutation rate, their atypical alleles seem to reflect an independent single mutation event that induces lost or decreased mutagenicity, thereby acting like a binary marker. Therefore, these findings indicate that the non-consensus alleles can be useful for achieving a further level of resolution within the binary Y-haplogroup tree.[27, 28]

Additionally, the 8-bp deletion mutation in the DYS385 flanking region and null alleles at DYS448 caused by the polymorphic 50f2C deletion (or non-b1/b3 class II) were also associated with haplogroups O3a2b-M7 and C3-M217, respectively. The 8-bp deletion mutation could be mapped to

Y-chromosome position 19261093-19261108 (DYS385a) or 19301828-19301843 (DYS385b) using human genome build NCBI36.3/hg18 and it has not been reported in dbSNP. The mutation only occurred in haplogroup O3a2b-M7 in these samples, so further analysis of Y-SNPs derived from M7 is needed to identify the relationship between the mutation and sublineages of haplogroup O3a2b-M7 in the Southeastern Asian population with notable frequencies of M7.[36, 65]

Based on this network analysis, the null alleles caused by the polymorphic 50f2C deletion seem to have originated from a single mutation event in haplogroup C3-M217. In the network-based cluster, two groups could be subdivided by Korean and Han Chinese males and by other Asians including Kyrgyzstani and Kalmyk. This finding suggests different evolutionary trajectories for each group. It is known that the deletion type of non-b1/b3 class II has risen to a high frequency in haplogroup C3-M217 (xC3a, C3c) in the Asian population.[53] Therefore, the deletion mutation may be also useful for revealing substructure within the haplogroup C3-M217, which occurs at a moderate frequency in Asia.

The Y-chromosomal haplogroup tree is consistently updated at a high level of detail (International Society of Genetic Genealogy, http://www.isogg.org/tree/). In the recent updated haplogroup O,[30] recently defined markers (KL2, JST002611, and P164) provide enhanced phylogenetic resolution of the Korean haplogroup O3a-M324 like those of Chinese Han.[30]

Although the absence of the JST002611 mutation has been reported in the Korean population,[32] a considerable number of samples were found to have that mutation in these surveyed Korean males (9.8%). The fact that this mutation is observed frequently in East Asia, including Japan, China, and mainland Southeast Asia (up to 21%),[24, 30, 65] supports the reliability of this data, and the presence of that mutation was also confirmed by sequence analysis.

Interestingly, it was possible hierarchically infer the non-equivalence of M117 and M133 from the presence of M117+/M133- samples. The network analysis seems to suggest a possible founder lineage by the ancestral state of M133 out of haplogroup O3a2c1a (M117+, M133+). However, the possibility of restoring the M133 mutation, e.g., back-mutation, is much more restricted due to the character of the 1-bp deletion mutation. Therefore, allocation of the M133 marker within phylogeny should be considered again and replaced with the more reliable binary marker M117 for designating haplogroup O3a2c1a, although further analysis is needed.

Haplogroups are determined by the pattern of Y-SNPs, which can also be tested and determined directly. However, the process of determining the haplogroup by direct testing of the Y-SNPs typically requires additional cost, time or considerable amounts of sample DNA. Therefore, there is considerable interest in predicting the haplogroup from a set of STR markers because many STRs are routinely genotyped in multiplex assays in forensic fields. Many methods for haplogroup prediction such as Y-STR allele-frequency and

machine-learning approaches have been developed[21, 22] based on a computational perspective. However, in this study, the haplotype status of DYS392, DYS393, DYS437, DYS438, DYS448 and DYS388 could be used for rough major haplogroup prediction through matching of the haplotypes with known haplogroups in the reference database, thereby enabling the inference of the geographic or ethnic origin of unknown samples. Additionally, several atypical alleles with a shared common sequence structure, for example the 8-bp deletion mutation at the DYS385 flanking region and the polymorphic 50f2C deletion reflect a single haplogroup background such that the haplogroup status could be inferred from the Y-STR variants.

In order to type the Y-chromosomal haplogroups according to the revised phylogenetic tree, specifically reflecting the updated phylogeny of haplogroup O, two different techniques, SBE reactions and AS-PCR assays were developed. SBE reactions using SNaPshot minisequencing were the most favored methods for SNP typing in forensic application,[66] which consists of the single base extension of an unlabeled primer that anneals one base upstream to the relevant SNP with fluorescence-labeled dideoxynucleotide (ddNTP). Allele designation is then possible by separating extended products and detecting the fluorescence by capillary electrophoresis. This SNP typing methodology proved to be robust, reliable and is extremely sensitive in various studies, especially applicable to degraded DNA analysis due to the possible reduction of PCR target sizes.[12, 13] However, the two steps of purification and the use of four fluorescence-labeled

ddNTP make the SBE reaction complicated and hard to interpret. Another technique, AS-PCR assay, is a one-step procedure that requires only a single PCR amplification with different length allele-specific primers, thereby allowing rapid analysis and ease data in interpretation. This AS-PCR assay can, however, often be tedious and time-consuming to establish, requiring lengthy optimization procedures. Moreover, the sensitivity or efficiency of the AS-PCR assay has not yet been evaluated for forensic samples. Therefore, six multiplex PCRs followed by SBE reaction were designed to have amplicon sizes ranging from 70 to 100 bp for analysis of highly degraded DNA and two multiplex AS-PCR assays were designed for a simple procedure and rapid analysis. The newly developed multiplexes were then validated.

The multiplex PCR systems followed by SBE reaction showed that the amount of DNA as small as 62pg and highly degraded DNA from old skeletal remains can be reliably typed with increasing PCR cycles. By increasing the PCR cycle number, however, unwanted phenomena such as allele dropout, heterozygous peak imbalance, increase in stutter peaks, or pull-up peaks could increase.[39] Fortunately, however, Y-SNP has no heterozygous allele and stutter peak associated with each allele, thus the unwanted effect of increased PCR cycle number can be avoided. On the other hand, multiplex AS-PCR assays were not adequate for applying increased PCR cycles to enhance sensitivity because non-specific amplifications can be observed. Instead, the multiplex AS-PCR assays showed a simple haplogroup designation like the general

forensic STR typing method. Collectively, validation tests revealed that multiplex PCR systems followed by SBE reaction were more suitable for analyzing low template and highly degraded DNA, whereas the multiplex AS-PCR assays were optimized for simple, rapid and reliable detection of intact or reference DNA, e.g. analysis of a large number of reference samples. Thus, the newly developed multiplex systems in this study are expected to enable researchers to work with more ease and efficiency when dealing with degraded DNA or a large number of DNA samples. However, it is important to be attentive to multiple peaks when interpreting mixed samples or drop-ins due to DNA contaminants.

V. Conclusion

Six multiplex PCRs followed by SBE reaction and two multiplex allele-specific PCR assays were developed to identify the most frequent East Asian Y-chromosomal haplogroups, especially with in-depth resolution for haplogroup O-M175. Sensitivity and efficiency tests revealed that multiplex PCR systems followed by SBE reaction were more suitable for analyzing low template and highly degraded DNA than multiplex allele-specific PCR assays. On the other hand, allele-specific PCR assays were useful for simple and rapid identification of haplogroups in a large number of samples. These developed multiplex sets will be useful tools for Y-chromosomal haplogroup determination

in degraded samples and forensic studies of the East Asian population. A total of 21 different haplogroups were identified by 33 Y-SNPs using multiplex PCRs followed by SBE reactions and multiplex allele-specific PCR assays, with further use of additional monoplex PCRs followed by SBE reaction in 1006 Korean males. When genotyping the SNPs, phylogenetic nonequivalence was found between SNPs M117 and M133, suggesting that the position of the M133 marker should be corrected. The additions of DYS446, DYS458, DYS635 and DYS449, which have high mutation rates, to the haplotypes increased the discriminatory power and the combined haplotype of 15 loci, DYS447, DYS458, DYS635, GATA H4.1 and the SWGDAM Y-STRs, was comparable to the haplotype of 17 loci in the AmpFlSTR® Yfiler™ kit. Whereas, the haplotype status of DYS392, DYS393, DYS437, DYS438, DYS448 and DYS388, which show slow mutation rates, could be used for a rough major haplogroup prediction through matching of the haplotypes with known haplogroups in the reference database. However, atypical alleles with shared a common sequence structure at the highly mutating Y-STRs, the 8-bp deletion mutation at DYS385 flanking region and the polymorphic 50f2C deletion, reflect a single haplogroup background such that the haplogroup status could be inferred from the Y-STR variants like Y-SNP markers. These findings and Y-STR/Y-SNP dataset provide useful information for building a Korean database and for elucidating Korean haplogroup substructures and resolving male genealogies in Asia.

Reference

1. Jobling MA, Pandya A, Tyler-Smith C. The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 1997;110:118-24.
2. Jobling MA, Tyler-Smith C. New uses for new haplotypes the human Y chromosome, disease and selection. Trends Genet 2000;16:356-62.
3. Jobling MA. In the name of the father: surnames and genetics. Trends Genet 2001;17:353-7.
4. Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, et al. Y chromosome sequence variation and the history of human populations. Nat Genet 2000;26:358-61.
5. Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 2003;4:598-612.
6. Roewer L, Croucher PJ, Willuweit S, Lu TT, Kayser M, Lessig R, et al. Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet 2005;116:279-91.
7. Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, et al. Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related populations. Hum Mol Genet 1997;5:1029-33.
8. Butler JM. Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis. Forensic Sci Rev 2003;15:91–111.
9. Jobling MA. Y-chromosomal SNP haplotype diversity in forensic analysis. Forensic Sci Int 2001;118:158-62.
10. Lessig R, Zoledziewska M, Fahr K, Edelmann J, Kostrzewa M, Dobosz T, et al. Y-SNP-genotyping - a new approach in forensic analysis. Forensic Sci Int

2005;154:128-36.

11. Dixon LA, Murray CM, Archer EJ, Dobbins AE, Koumi P, Gill P. Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes. Forensic Sci Int 2005;154:62-77.
12. Bouakaze C, Keyser C, Amory S, Crubézy E, Ludes B. First successful assay of Y-SNP typing by SNaPshot minisequencing on ancient DNA. Int J Legal Med 2007;121:493-9.
13. Amigo J, Phillips C, Lareu M, Carracedo A. The SNPforID browser: an online tool for query and display of frequency data from the SNPforID project. Int J Legal Med 2008;122:435-40.
14. Prinz M, Sansone M. Y chromosome-specific short tandem repeats in forensic casework. Croat Med J 2001;42:288-91.
15. Yoshida Y, Fujita Y, Kubo S. Forensic casework of personal identification using a mixture of body fluids from more than one person by Y-STRs analysis. J Med Invest 2004;51:238-42.
16. Willuweit S, Roewer L; International Forensic Y Chromosome User Group. Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int Genet 2007;1:83-7.
17. Mršić G, Gršković B, Vrdoljak A, Popović M, Valpotić I, Anđelinović Š, et al. Croatian national reference Y-STR haplotype database. Mol Biol Rep 2012;39:7727-41.
18. Taylor D, Nagle N, Ballantyne KN, van Oorschot RA, Wilcox S, Henry J, et al. An investigation of admixture in an Australian Aboriginal Y-chromosome STR database. Forensic Sci Int Genet 2012;6:532-8.

19. Bosch E, Calafell F, Santos FR, Pérez-Lezaun A, Comas D, Benchemsi N, et al. Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. Am J Hum Genet 1999);65:1623-38.
20. Ballantyne KN, Keerl V, Wollstein A, Choi Y, Zuniga SB, Ralf A, et al. A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. Forensic Sci Int Genet 2012;6:208-18.
21. Schlecht J, Kaplan ME, Barnard K, Karafet T, Hammer MF, Merchant NC. Machine-learning approaches for classifying haplogroup from Y chromosome STR data. PLoS Comput Biol 2008;4:e1000093.
22. Athey TW. Haplogroup prediction from Y-STR values using an allele-frequency approach. J Genet Geneal 2005;1:1-7.
23. Onofri V, Alessandrini F, Turchi C, Fraternale B, Buscemi L, Pesaresi M, et al. Y-chromosome genetic structure in sub-Apennine populations of Central Italy by SNP and STR analysis. Int J Legal Med 2007;121:234-7.
24. Nonaka I, Minaguchi K, Takezaki N. Y-chromosomal binary haplogroups in the Japanese population and their relationship to 16 Y-STR polymorphisms. Ann Hum Genet 2007;71:480-95.
25. Robino C, Crobu F, Di Gaetano C, Bekada A, Benhamamouch S, Cerutti N, et al. Analysis of Y-chromosomal SNP haplogroups and STR haplotypes in an Algerian population sample. Int J Legal Med 2008;122:251-5.
26. Kayser M, Brauer S, Weiss G, Schiefenhövel W, Underhill PA, Stoneking M. Independent histories of human Y chromosomes from Melanesia and Australia. Am J Hum Genet 2001;68:173-90.

27. Myres NM, Ekins JE, Lin AA, Cavalli-Sforza LL, Woodward SR, Underhill PA. Y-chromosome short tandem repeat DYS458.2 non-consensus alleles occur independently in both binary haplogroups J1-M267 and R1b3-M405. Croat Med J 2007;48:450-9.
28. Myres NM, Ritchie KH, Lin AA, Hughes RH, Woodward SR, Underhill PA. Y-chromosome short tandem repeat intermediate variant alleles DYS392.2, DYS449.2, and DYS385.2 delineate new phylogenetic substructure in human Y-chromosome haplogroup tree. Croat Med J 2009;50:239-349.
29. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res 2008;18:830-8.
30. Yan S, Wang CC, Li H, Li SL, Jin L; Genographic Consortium. An updated tree of Y-chromosome Haplogroup O and revised phylogenetic positions of mutations P164 and PK4. Eur J Hum Genet 2011;19:1013-5.
31. Kim SH, Han MS, Kim W, Kim W. Y chromosome homogeneity in the Korean population. Int J Legal Med 2010;124:653-7.
32. Kim SH, Kim KC, Shin DJ, Jin HJ, Kwak KD, Han MS, et al. High frequencies of Y-chromosome haplogroup O2b-SRY465 lineages in Korea: a genetic perspective on the peopling of Korea. Investig Genet 2011;2:10.
33. Park MJ, Lee HY, Yoo JE, Chung U, Lee SY, Shin KJ. Forensic evaluation and haplotypes of 19 Y-chromosomal STR loci in Koreans. Forensic Sci Int 2005;152:133-47.
34. Park MJ, Lee HY, Kim NY, Sim JE, Yang WI, Cho SH, et al. Genetic characteristics of 22 Y-STR loci in Koreans. Korean J Leg Med 2007;31:162-70.

35. Lee HY, Park MJ, Kim NY, Sim JE, Yang WI, Shin KJ. Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. Forensic Sci Int Genet 2010;4:275-80.
36. Shi H, Dong YL, Wen B, Xiao CJ, Underhill PA, Shen PD, et al. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. Am J Hum Genet 2005;77:408-19.
37. Xue Y, Zerjal T, Bao W, Zhu S, Shu Q, Xu J, et al. Male demography in East Asia: a north-south contrast in human population expansion times. Genetics 2006;172:2431-9.
38. Zhong H, Shi H, Qi XB, Xiao CJ, Jin L, Ma RZ, et al. Global distribution of Y-chromosome haplogroup C reveals the prehistoric migration routes of African exodus and early settlement in East Asia. J Hum Genet 2010;55:428-35.
39. Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. Forensic Sci Int 2000;112:17-40.
40. Cowen S, Debenham P, Dixon A, Kutranov S, Thomson J, Way K. An investigation of the robustness of the consensus method of interpreting low-template DNA profiles. Forensic Sci Int Genet 2011;5:400-6.
41. Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, et al. DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Int J Legal Med 2006;120:191-200.
42. Lin YW, Hsu CL, Yen PH. A two-step protocol for the detection of rearrangements at the AZFc region on the human Y chromosome. Mol Hum Reprod

2006;12:347-51

43. Excoffier L, Lischer, HE L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 2010;10:564-7
44. Gomes V, Sánchez-Diz P, Amorim A, Carracedo A, Gusmão L. Digging deeper into East African human Y chromosome lineages. Hum Genet 2010;127:603-13
45. Wiegand P, Kleiber M. Less is more--length reduction of STR amplicons using redesigned primers. Int J Legal Med 2001;114:285-7.
46. Butler JM, Shen Y, McCord BR. The development of reduced size STR amplicons as tools for analysis of degraded DNA. J Forensic Sci 2003;48:1054-64.
47. Kainz P. The PCR plateau phase - towards an understanding of its limitations. Biochim Biophys Acta 2000;1494:23-7.
48. Walsh PS, Erlich HA, Higuchi R. Preferential PCR amplification of alleles: mechanisms and solutions. PCR Methods Appl 1992;1:241-50.
49. Lee HY, Park MJ, Chung U, Lee HY, Yang WI, Cho SH, et al. Haplotypes and mutation analysis of 22 Y-chromosomal STRs in Korean father-son pairs. Int J Legal Med 2007;121:128-35.
50. Hanson EK, Ballantyne J. Comprehensive annotated STR physical map of the human Y chromosome: Forensic implications. Leg Med (Tokyo) 2006;8:110-20.
51. Fernandes S, Huellen K, Goncalves J, Dukal H, Zeisler J, Rajpert De Meyts E, et al. High frequency of DAZ1/DAZ2 gene deletions in patients with severe oligozoospermia. Mol Hum Reprod 2002;8:286-98.
52. Kuroda-Kawaguchi T, Skaletsky H, Brown LG, Minx PJ, Cordum HS, Waterston RH, et al. The AZFc region of the Y chromosome features massive palindromes and

uniform recurrent deletions in infertile men. Nat Genet 2001;29:279-86.

53. Balaresque P, Bowden GR, Parkin EJ, Omran GA, Heyer E, Quintana-Murci L, et al. Dynamic nature of the proximal AZFc region of the human Y chromosome: multiple independent deletion and duplication events revealed by microsatellite analysis. Hum Mutat 2008;29:1171-80.
54. Füredi S, Woller J, Pádár Z, Angyal M. Y-STR haplotyping in two Hungarian populations. Int J Legal Med 1999;113:38-42.
55. King TE, Bosch E, Adams SM, Parkin EJ, Rosser ZH, Jobling MA. Inadvertent diagnosis of male infertility through genealogical DNA testing. J Med Genet 2005;42:366-8.
56. Kayser M, Kittler R, Erler A, Hedman M, Lee AC, Mohyuddin A, et al. A comprehensive survey of human Y-chromosomal microsatellites. Am J Hum Genet 2004;74:1183-97.
57. Mizuno N, Nakahara H, Sekiguchi K, Yoshida K, Nakano M, Kasai K. 16 Y chromosomal STR haplotypes in Japanese. Forensic Sci Int 2008;174:71-6.
58. Parkin EJ, Kraayenbrink T, Opgenort JR, van Driem GL, Tuladhar NM, de Knijff P, et al. Diversity of 26-locus Y-STR haplotypes in a Nepalese population sample: Isolation and drift in the Himalayas. Forensic Sci Int 2006;166:176-81.
59. Chang YM, Perumal R, Keat PY, Kuehn DL. Haplotype diversity of 16 Y-chromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia. Forensic Sci Int 2006;167:70-6.
60. Roewer L, Krüger C, Willuweit S, Nagy M, Rodig H, Kokshunova L, et al. Y-chromosomal STR haplotypes in Kalmyk population samples. Forensic Sci Int 2007;173:204-9.

61. Gutiérrez-Alarcón AB, Moguel-Torres M, León-Jiménez AK, Cuéllar-Nevárez GE, Rangel-Villalobos H. Allele and haplotype distribution for 16 Y-STRs (AmpFlSTR Y-filer kit) in the state of Chihuahua at North Center of Mexico. Leg Med (Tokyo) 2007;9:154-7.
62. Sánchez C, Barrot C, Xifró A, Ortega M, de Aranda IG, Huguet E, et al. Haplotype frequencies of 16 Y-chromosome STR loci in the Barcelona metropolitan area population using Y-Filertrade mark kit. Forensic Sci Int 2007;172:211-7.
63. Gusmão L, Krawczak M, Sánchez-Diz P, Alves C, Lopes A, Beleza S, et al. Bimodal allele frequency distribution at Y-STR loci DYS392 and DYS438: no evidence for a deviation from the stepwise mutation model. Int J Legal Med 2003;117:287-90.
64. Cinnioğlu C, King R, Kivisild T, Kalfoğlu E, Atasoy S, Cavalleri GL, et al. Excavating Y-chromosome haplotype strata in Anatolia. Hum Genet 2004;114:127-48.
65. Karafet TM, Hallmark B, Cox MP, Sudoyo H, Downey S, Lansing JS, et al. Major east-west division underlies Y chromosome stratification across Indonesia. Mol Biol Evol 2010;27:1833-44.
66. Sobrino B, Brión M, Carracedo A. SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci Int 2005;154:181-94.

< ABSTRACT(IN KOREAN)>

Understanding the Y chromosome variation by haplogroup and haplotype analyses in a Korean population

<지도교수 신경진>

연세대학교 대학원 의과학과

박명진

Y 염색체에서 재조합이 일어나지 않는 영역은 생물학적 증거물의 부계 조상이나 유래된 지역에 관한 정보를 제공할 뿐만 아니라 서로 다른 부계혈통을 구분하는데 쓰인다. 본 연구에서는 Y 염색체의 단일염기다양성(SNP)와 짧은염기반복서열(STR) 마커를 이용하여 한국인 집단의 Y 염색체 유전적 구조에 관한 특성을 파악하고자 한다. 이를 위하여 SBE (single base extension) 반응을 이용한 6개의 다중연쇄중합반응(multiplex)과 대립유전자 특이적 증폭(allele-specific PCR)을 이용한 2개의 다중연쇄중합반응을 한국인 뿐만 아니라 동아시아에서 빈번하게 나타나는 하플로그룹을 결정할 수 있도록 개발하였다. 이러한 개발된 다중연쇄중합반응 세트를 평가한 결과 SBE 반응을 이용한 다중연쇄중합반응은 적은 양의 DNA와 매우 분해된 DNA 분석에 보다 효과적인 반면에 대립유전자 특이적 증폭을 이용한 다중연쇄중합반응은 분석에 보다 빠르고 간편한 방법으로 많은 양의 시료를 분석하는데 유용할 것으로 생각된다. 다중연쇄중합반응을 이용하여 하플로그룹 O에서 새롭게 지정된 PK4, KL2와 P164 마커를 포함한 33개의 Y 단일염기다양성 마커로 한국인 집단을 분석하였고 그 결과 총 21개의 서로 다른 하플로그룹을 결정할 수 있었다. 이때 하플로그룹 O3a2c1a를 결정한다고 알려진 M117과 M133 마커가 일부 동일한 시료에서 M117 마커는 변이형 M133마커는 와일드형으로 나타나 M117과 M133 마커가 하플로그룹을

결정하는 계층적인 tree에서 같은 위치가 아니고 M133마커의 위치가 수정되어야 할 것으로 생각된다. 짧은염기반복서열 마커와 그 일배채형 (haplotype)을 하플로그룹과 관련지어 분석하였을 때 상대적으로 다른 짧은염기반복서열 마커보다 돌연변이율이 낮은 DYS392, DYS393, DYS437, DYS438, DYS448과 DYS388 마커와 이들 마커로 이루어진 일배채형이 보다 계통유전적인 정보를 가지고 있어 이들 마커의 정보로부터 어느 정도 하플로그룹을 예측할 수 있었다. 반면에 DYS449나 DYS458 마커 같은 상대적으로 돌연변이율이 높은 마커들은 서로 다른 부계혈통을 구분하는데 보다 더 유용하였다. 하지만 상대적으로 돌연변이율이 높은 DYS447, DYS449, DYS458과 DYS464 마커에서 일관된 염기서열구조를 가진 비전형적 짧은 대립유전자와 비전형 중간대립유전자는 특정 하플로그룹 내 서브구조를 설명할 수 있었다. 게다가 DYS385의 반복서열주변의 소실 돌연변이(deletion mutation)와 DYS448의 null 대립유전자들도 단일 하플로그룹 배경과 연관되어 있었다. 따라서 이러한 변이대립유전자는 binary 마커와 같이 그 변이형이 single mutation event에 의해 유래된 것이라는 가설을 밑받침 한다. 분해능이 높은 하플로그룹과 이와 연계된 짧은염기반복서열 마커의 일배체형 정보는 한국인 집단의 구조를 파악할 수 있게 할 뿐만 아니라 하플로그룹 배경이나 공통부계조상을 추정하는데 도움을 제공할 것이다.

핵심되는 말 : Y 염색체, 하플로그룹, 단일염기다양성, 짧은염기반복서열, 비전형적인 대립유전자, 한국인

PUBLICATION LIST

1. Park MJ, Lee HY, Kim NY, Lee EY, Yang WI, Shin KJ. Y-SNP miniplexes for East Asian Y-chromosomal haplogroup determination in degraded DNA. Forensic Sci Int Genet 2013;7:75-81.
2. Park MJ, Lee HY, Yang WI, Shin KJ. Understanding the Y chromosome variation in Korea - relevance of combined haplogroup and haplotype analyses. Int J Legal Med 2012;126:588-99.
3. Lee HY, Park MJ, Kim NY, Sim JE, Yang WI, Shin KJ. Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. Forensic Sci Int Genet 2010;4:275-80.
4. Park MJ, Shin KJ, Kim NY, Yang WI, Cho SH, Lee HY. Characterization of deletions in the DYS385 flanking region and null alleles associated with AZFc microdeletions in Koreans. J Forensic Sci 2008;53:331-4.
5. Park MJ, Lee HY, Kim NY, Sim JE, Yang WI, Cho SH, Shin KJ. Genetic characteristics of 22 Y-STR loci in Koreans. Korean J Leg Med 2007;31:162-70.
6. Lee HY, Park MJ, Chung U, Lee HY, Yang WI, Cho SH, Shin KJ. Haplotypes and mutation analysis of 22 Y-chromosomal STRs in Korean father-son pairs. Int J Legal Med 2007;121:128-35.
7. Park MJ, Lee HY, Yoo JE, Chung U, Lee SY, Shin KJ. Forensic evaluation and haplotypes of 19 Y-chromosomal STR loci in Koreans. Forensic Sci Int. 2005;152:133-47.